Characterisation of Endogenous KRAB Zinc Finger Proteins

Catherine Crawford

PhD Thesis

University of Edinburgh

2009

Declaration

I declare that the work presented in this PhD thesis is original research carried out by me, and that where the work of others has been presented, it has been stated accordingly. This work has not been submitted for any other degree.

Catherine Crawford, August 2009

Acknowledgements

There's so many people that I'd like to thank, that have been supportive of me in the past 3 and a bit years, either in the lab or over the phone or in the pub, I hope I manage to remember everyone. Sorry if I don't…

Well, first of all I'd like to thank those people in the Unit. Many thanks to Wendy and Heidi, my two supervisors, who helped me with lab work for the first three years, and who were really great (and speedy) at reading and correcting my thesis. Thanks also to Wendy for letting me go to Cold Spring Harbor, which was one of the highlights of my PhD. The HGU, and E3 in particular, has been a cracking place to work, and I've made some fantastic friends there. Thanks to everyone in the Bickmore lab and many others for their help and suggestions over the years, in particular Ragnhild, Nick, Celine and Duncan. Many thanks also go to David and Colin for their help with all the ubiquitin-SUMO work. Loads of the students have aided me in all matters alcohol-related, and for this I am grateful, especially to Andy, Jon and Leeanne. Extra special thanks go to the ladies (Shelagh, Liz, Jenny, Diana, Lorna, and I guess P^2 as an honorary gentleman), who have been my tea-drinking, knitting, crafting, cake-eating, crossword-doing, gossiping buddies, and the affectionately titled 'langers' (Donncha, Jamie, Luke, Ross) for the occasional trip to the Telford, and never failing to entertain me, usually by lowering to chat level to grade C. Lots of love also goes to Rachel and Jen, my partners in champagne-drinking-Po-Na-Na-90s-music-dancing crime. Oh, and I promised I would namecheck the Midori club, and now I have, thanks to them for one of the drunkest nights of my life. My friends outside of work have helped me stay a little bit sane, and so Craig, Helen and the uni lot, especially Tom and Rachel, deserve a mention.

To my family: thanks for all your support, love and advice. Mum and Julia, you have been amazing, thank you so much, and I'm sorry I've not visited as much as I should have recently! Maybe now it's finished I'll bore you about it.

Finally, and most importantly, John: I couldn't have done this without you. Thank you for helping me when it was bad (making my dinner, looking after me when I was moody, putting up with my late nights, either in the lab or in the pub) and making everything that was good even better. I love you chicken.

Abstract

The Krüppel-associated box (KRAB) zinc finger protein (ZFP) genes comprise one of the largest gene families in the mammalian genome, encoding transcription factors with an N-terminal KRAB domain and C-terminal zinc fingers. The KRAB domain interacts with a co-repressor protein, KAP-1, which can recruit various factors causing transcriptional repression of genes to which KRAB ZFPs bind. Little is currently known about the gene targets of the ~400 human and mouse KRAB ZFPs.

Many KRAB ZFPs interact with factors other than KAP-1. To identify proteins that may interact with one particular KRAB ZFP, Zfp647, I previously carried out a yeast two-hybrid screen using the full-length Zfp647 sequence and a mouse embryonic cDNA library. I have now tested the interactions from this screen for their specificity for Zfp647. I show that Zfp647 can interact with itself and at least 20 other KRAB ZFPs through their zinc finger domains, and have confirmed the Zfp647 self-interaction by *in vitro* co-immunoprecipitation. In my yeast two-hybrid screen, Zfp647 bound to KAP-1 as well as another related protein, ARD1/Trim23. Zfp647 also interacts with proteins that function in ubiquitylation. I have found evidence to suggest that Zfp647 may also interact with proteins encoding jumonji domains both by yeast two-hybrid assay and by co-immunoprecipitation from NIH/3T3 cell extracts. We have previously found that Zfp647 localises to non-heterochromatic nuclear foci in differentiated ES cells, which also contain KAP-1 and HP1, and which lie adjacent to PML nuclear bodies in a high proportion of cells. I have found that these foci are also visible in pMEFs, but not NIH/3T3 tissue culture cells. Immunofluorescence studies with antibodies against proteins from the yeast two-hybrid screen have not shown any significant co-localisation with Zfp647.

KAP-1 is sumoylated *ex vivo*, as are two human KRAB ZFPs. Because Zfp647 lies adjacent to PML nuclear bodies and can associate with proteins involved in post-translational modification, I tested whether Zfp647 is also modified. I characterised a

sheep _-Zfp647 antibody previously created in the lab and have shown that it detects Zfp647 by western blot, but not by immunofluorescence. I show that treatment of NIH/3T3 cells with NEM, which prevents the removal of protein modifications, leads to the appearance of higher molecular weight forms of Zfp647. Modification of Zfp647 is not dependent on KAP-1, which is known to function as a SUMO E3 ligase. Attempts to classify the modification as either ubiquitin, SUMO or NEDD8 have suggested that Zfp647 may be mono-ubquitylated. The larger modified forms of Zfp647 are present in both NIH/3T3 and ES cells. Interestingly, I found that the modification profile of the protein changes over the course of ES cell differentiation, during which time Zfp647 relocalises to punctate nuclear foci; thus Zfp647 modification may be involved in this process.

Table of Contents

Chapter 2: Materials and Methods

Chapter 3: Protein-Protein Interactions of Zfp647

Chapter 4: Post-Translational Modification of Zfp647

List of Figures

Chapter 1

Chapter 3

Chapter 4

List of Tables

Chapter 1

Chapter 2

Chapter 3

List of Abbreviations

A	adenine
A_x	absorbance of light at x nm
ADP	adenosine diphosphate
APL	acute promyelocyte leukaemia
ALT	alternative lengthening of telomeres
APB	ALT-associated PML-NB
ATPase	adenosine triphosphate hydrolase
bp	basepair(s)
C	cytosine
C-terminal	carboxy-terminal
cDNA	complementary DNA
ChIP	chromatin immunoprecipitation
CpG	cytosine-phosphate-guanine
D-MEM	Dulbecco's-minimal Eagle's medium
DAPI	4',6-diamidino-2-phenylindole
DMP	dimethyl pimelimidate
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
dNTP	deoxynucleotide triphosphate
DSB	double strand break
DTT	dithiothreitol
EC	embryonic carcinoma
EDTA	ethylene diamine tetraacetic acid
EGTA	ethylene glycol tetraacetic acid
ES	embryonic stem
FCS	foetal calf serum
G	guanine
G-MEM	Glasgow's minimal Eagle's medium
GFP	green fluorescent protein
GST	glutathione-S-transferase
HA	haemagglutinin
HAT	histone acetyltransferase
HDAC	histone deacetylase
His_6	hexa-histidine
HMTase	histone methyltransferase
HR	homologous recombination
HRP	horseradish peroxidase
HSA	*Homo sapiens* chromosome

HP1	heterochromatin protein-1
IAA	isoamylalcohol
Ig	immunoglobulin
INM	inner nuclear membrane
IP	immunoprecipitation
IPTG	isopropyl _-D-1-thiogalactopyranoside
KAP-1	KRAB-associated protein-1
kDa	kilo Dalton(s)
KRAB	Krüppel-associated box
KRIP-1	KRAB-interacting protein-1
LB	Luria Bertani broth
LCR	locus control region
LIF	leukaemia inhibitory factor
M	molar
Mb	megabasepair(s)
MEM	minimal Eagle's medium
MMU	*Mus musculus* chromosome
mRNA	messenger RNA
N-terminal	amino-terminal
ND10	nuclear domain 10
NEM	N-ethylmaleimide
NLS	nuclear localisation signal
nm	nanometre
NP-40	Nonidet P-40
NRL	nucleosome repeat length
OD_x	optical density at x nm wavelength
PAGE	polyacrylamide gel electrophoresis
PARP	poly(ADP-ribose) polymerase
PBS	phosphate-buffered saline
PBST	PBS containing 0.1% Tween-20
PCR	polymerase chain reaction
pFA	paraformaldehyde
PHD	plant homeodomain
pMEF	primary mouse embryonic fibroblast
PML	promyelocyte leukaemia
PML-NB	PML nuclear body
POD	PML oncogenic domain
PrE	Primitive endoderm
RA	retinoic acid
RB	retinoblastoma protein

RBAK	RB-associated KRAB repressor
RBCC	RING, b-boxes and coiled coil
RING	really interesting new gene
RNA	ribonucleic acid
SD	Sabouraud dextrose
SDS	sodium dodecyl sulphate
SIM	SUMO-interacting motif
SUMO	small ubiquitin-like modifier
T	thymine
TAE	Tris-EDTA acetic acid buffer
TBE	Tris-EDTA boric acid buffer
TE	Tris-EDTA
Tif1_	transcription intermediary factor-1_
TNT	*in vitro* transcription and translation
Trim	tripartite motif protein
TSS	transcription start site
UV	ultra-violet radiation
WBR	western blocking reagent
YPD	yeast peptone dextrose
ZAD	zinc finger-associated C4DM domain
ZNF	human zinc finger protein
Zfp	murine zinc finger protein

Amino Acid Code

Name	Three-letter name	Single-letter name
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartic acid	Asp	D
Cysteine	Cys	C
Glutamic acid	Glu	E
Glutamine	Gln	Q
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

Chapter 1: Introduction

1.1. Chromatin

1.1.1. Chromatin Structure

In eukaryotes, packaging of the genome into the nucleus is achieved by condensing DNA into tightly wound structures by proteins. This mixture of DNA and proteins, roughly a 1:2 ratio by mass (Lee and Young, 2000), is known as chromatin. At the basic level, 146bp of the DNA helix is wound around 8 highly conserved histone proteins (a tetramer composed of two H3 and two H4 histone dimers, flanked by two heterodimers of histones H2A and H2B) (Luger *et al*, 1997), giving a structure called a nucleosome. Individual nucleosomes are held together by the H1 linker histone proteins. The average number of H1 molecules per nucleosome can vary from 0.5 in mouse embryonic stem (ES) cells (Fan *et al*, 2005) to 1.3 in chicken erythrocytes (Bates and Thomas, 1981). The distance between each nucleosome, the nucleosome repeat length (NRL), is also variable, ranging from 189bp in mouse ES cells (Fan *et al*, 2005) to 212bp in chicken erythrocytes (Bates and Thomas, 1981). Histones themselves have a negative effect on transcription: DNA bound by the core histones undergoes transcription at 15-30% lower levels than naked DNA, and addition of linker histone H1 represses transcription even further (Ura *et al*, 1995).

The simplest chromatin fibre, consisting of DNA bound by nucleosomes, is known as the 10nm fibre, or informally as 'beads on a string'. The 10nm fibre can be packaged up into a helix 30nm in diameter (the 30nm fibre) (Figure 1.1). The exact structure of the 30nm fibre is still hotly debated, with some groups favouring a 'one-start' helix (Robinson *et al*, 2006; Routh *et al*, 2008), and others championing a 'two-start' structure (Dorigo *et al*, 2004). The 30nm fibre can be further condensed on many levels to eventually form a 700nm structure. This level of packaging is not seen throughout the cell cycle: the 700nm fibre is only visible during cell division as the classic metaphase x-shaped chromosomes, to ensure the tangle-free separation of sister chromatids or chromosomes.

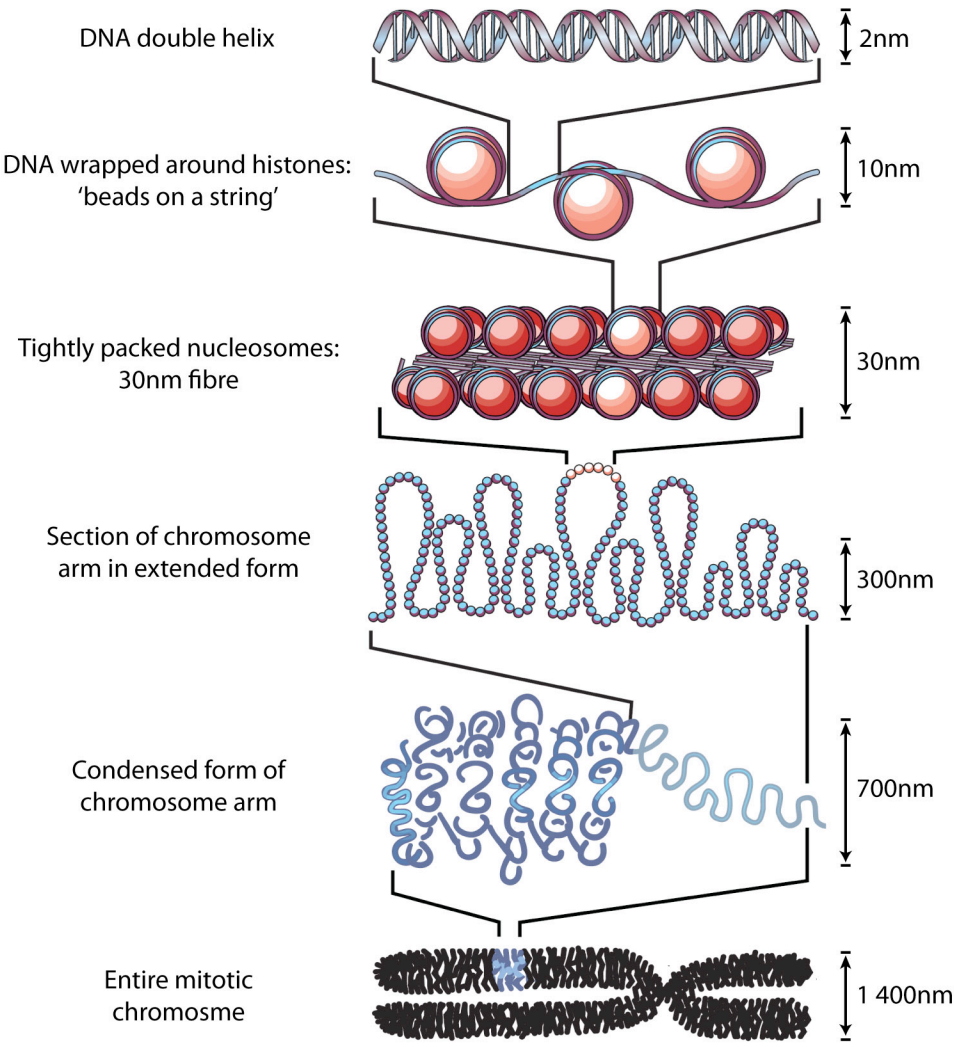


Figure 1.1. Levels of DNA compaction. Naked DNA is wrapped around histone octamers to form nucleosomes - the beads on a string conformation, or 10nm fibre. This chromatin fibre can be packaged into the 30nm fibre, which is further condensed to eventually give the classic x-shaped chromosomes seen at mitosis, measuring 1.4µm. Figure adapted from Felsenfeld and Groudine, 2003.

Given its typical folded structure, some degree of remodelling of the chromatin is required for the access of the transcription machinery and the synthesis of mRNA to occur. Chromatin is often labelled as being either 'open', i.e. with a relatively loose packaging arrangement, or 'closed', i.e. a tightly wound, condensed structure. Highly condensed, late replicating chromatin is known as heterochromatin, whereas early replicating chromatin with an open structure is known as euchromatin. The permanent, or constitutive, heterochromatin of a cell is composed of large arrays of tandem repeats often found at centromeric and pericentromeric regions, and therefore contains repetitive, relatively transcriptionally inactive sequences. The more transcriptionally active regions of the genome are euchromatic, and for this reason it was originally thought that chromatin structure was dependent on transcriptional activity (Felsenfeld and Groudine, 2003). However, work carried out on a more biophysical level has shown that gene density is a greater determinant of whether a genomic region is open or closed (Gilbert *et al*, 2004). The open structure of gene-rich domains would grant the access of transcription factors to the DNA, allowing transcription to occur relatively easily, whereas genes in a condensed, gene-poor region would require further factors to loosen the chromatin structure before the initiation of RNA synthesis (Sproul *et al*, 2005).

1.1.2. Chromatin Remodelling

Nucleosome positioning is important in determining the efficiency of transcription initiation and elongation. The transcription start site (TSS) of a gene is typically depleted of nucleosomes, with nucleosome free regions of ~150bp being found approximately 200bp upstream of start codons in yeast (Yuan *et* al, 2005). However, nucleosomes can still present a physical barrier to the elongating polymerase within the remainder of the coding sequence. The chromatin must therefore be rearranged in such a way to remove the histones from the path of the polymerase whilst maintaining the overall integrity of its structure. Thus it was shown that a purified human transcription system cannot efficiently transcribe a DNA template when chromatinised, but addition of certain proteins that can disrupt

nucleosome positioning on a stretch of DNA allows transcription to occur (Orphanides *et al*, 1998). Proteins that alter the pattern of nucleosomes are known as chromatin or nucleosome remodellers.

Eukaryotes possess at least 5 ATP-dependent chromatin remodeller families: SWI/SNF, ISWI, NuRD/Mi-2/CHD, INO80 and SWRI (Saha *et al*, 2006). Each chromatin remodelling complex contains at its heart an ATPase subunit, which is homologous to the SWI/SNF2 subfamily of the DEAD/H box family of nucleic acid-stimulated ATPases (Eisen *et al*, 1995). The two best studied families are the SWI/SNF and ISWI proteins. The ATPase domains in these complexes are identical; therefore the other components of the complex determine its overall function. The SWI/SNF remodellers are large multisubunit complexes that can adjust nucleosome position and the nucleosome/DNA structure, and that are capable of exchanging histone dimers or octamers (Martens and Winston, 2003). In contrast, the smaller ISWI complexes generate nucleosomes with altered positions but in standard structures, and remodel the chromatin to achieve regular spacing of nucleosomes, which promotes the organisation of the chromatin fibre into compact higher order structures (Deuring *et al*, 2000; Goldmark *et al*, 2000; reviewed in Racki and Narlikar, 2008).

1.1.3. Histone Modifications and their Effect on Transcription

Histone proteins have two domains: the globular C-terminal domain, which lies at the interior of the nucleosome; and the N-terminal (or C-terminal) 'tail' of the protein, which protrudes from the nucleosome (Luger *et al*, 1997). The tail region can interact with the DNA wrapped around the histone proteins, and with other chromatin-associated factors. Many proteins can modify histone tails, adding acetyl, methyl or ubiquitin groups for example (see Figure 1.2). In 2000, Strahl and Allis published their theory of the histone code: that histone modifications form a code that can be 'read' by proteins capable of recognising specific modifications. These

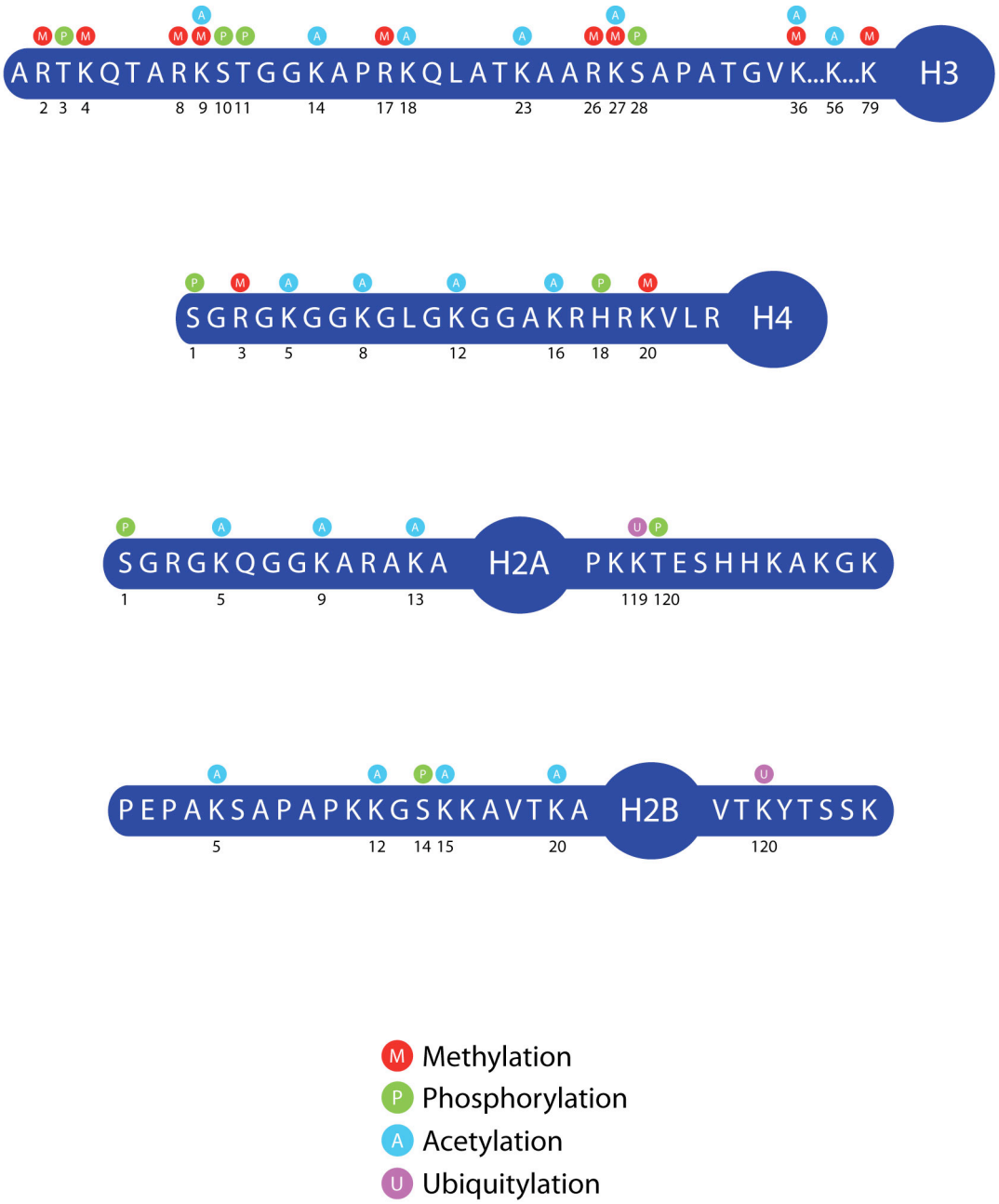


Figure 1.2. Common modifications of the four nuclesomal histones. The long long stretch of amino acids depict the N- (and C-) terminal tails of the histones; the circular region represents the core domain of the protein. Red circles denote methylation, green circles denote phosphorylation, blue circles denote acetylation and purple circles denote ubiquitylation. Adapted from Kouzarides, 2007.

proteins may add other marks to proximal histone tails, or alter the surrounding chromatin structure, making it more or less conducive to transcription (Strahl and Allis, 2000). Particular marks are therefore associated with certain transcriptional states, and we would be able to tell what is happening at a gene by the modifications present on the histones at that sequence. Antibodies highly specific for histone modifications can be used in chromatin immunoprecipitation (ChIP) to ascertain the modifications present at a DNA sequence, and this is now commonly combined with genomic microarrays (ChIP-chip) or high-throughput sequencing (ChIP-seq) to generate genome-wide maps of histone modifications in different cell types (Barski *et al*, 2007; Mikkelsen *et al*, 2007).

1.1.3.1. Histone Acetylation

Histones are rich in the positively charged amino acid lysine, which aids in their interaction with negatively charged DNA. However, lysine residues are commonly modified by the acetyl modification. The acetyl group is negatively charged, which neutralises the lysine residues and relaxes the interaction between DNA and histones (Hong *et al*, 1993). This makes the DNA more accessible to transcription factors and RNA polymerases, increasing the rate of transcription of the acetylated sequence; thus histone acetylation is linked to expressed genes (Hebbes *et al*, 1988). Acetylation of histones H3 and H4 peaks at the TSSs of active genes in humans, mice and yeast (Pokholok *et al*, 2005; Bernstein *et al*, 2005). Tse and colleagues have also shown that hyperacetylation of histone tails *in vitro* disrupts higher-order chromatin folding, and is accompanied by a 15-fold increase in transcription of the template tested (Tse *et al*, 1998), indicating that histone acetylation may also destabilise nucleosome-nucleosome interactions.

Histones H2A, H2B, H3 and H4 are all acetylated on their N-terminal tails (Figure 1.2). The addition of this modification is carried out by histone acetyltransferase (HAT) enzymes such as the functionally related proteins p300 and Creb binding-protein (p300/CBP), and p300/CBP-associated factor (PCAF). These

proteins are typically found as part of large, multiprotein complexes. Histone modifications are reversible: acetyl groups are removed from histones by histone deacetylase (HDAC) enzymes, which, like HATs, are typically found as subunits of large complexes. As expected, recruitment of HDAC-containing complexes to a gene results in the downregulation of its transcription. The two major HDAC complexes are NuRD (nucleosome remodelling and histone deacetylation) and Sin3 (reviewed by Cunliffe, 2008). These complexes share four subunits: the histone deacetylases HDAC1 and HDAC2, and the histone binding proteins RbAp46 and RbAp48. NuRD additionally has ATP-dependent nucleosome remodelling activity, through the Mi-2_ and Mi-2_ proteins (Denslow and Wade, 2008), thus this complex is capable of affecting both histone modification and nucleosome organisation.

1.1.3.2. Histone Methylation

Histone proteins can also be modified by the addition of methyl groups onto a lysine or arginine residue (Figure 1.2). Histones are mono-, di- or tri-methylated, which can be recognised as distinct marks by other proteins, giving an extra layer of complexity to this modification. The first histone methyltransferases (HMTases) to be discovered were the suppressor of variegation (Suv39h1 and Suv39h2) proteins (Rea *et al*, 2000; O'Carroll, 2000). These enzymes encode the highly conserved SET domain, named after the proteins it was first discovered in: Suppressor of Variegation 3-9, Enhancer of Zeste and trithorax (Jones and Gelbart, 1993; Tschiersch *et al*, 1994). The SET domain was subsequently identified in a number of other HMTases including G9a and SETDB1 (Tachibana *et al*, 2001; Schultz *et al*, 2002). Mammalian Suv39h1 and Suv39h2 tri-methylate H3K9 (Tachibana *et al*, 2001), and SETDB1 is also known to selectively methylate H3K9 (Schultz *et al*, 2002). The H3K9me mark is typically found at the TSSs of genes in repressed chromatin: H3K9me3 is a characteristic of heterochromatin, whereas H3K9me1 and H3K9me2 are enriched within silent euchromatic regions (Rice *et al*, 2003).

The tri-methylation of H3K27 is catalysed by the polycomb group repressive complex 2 (PRC2) subunits Enhancer of Zeste 2 (EZH2) (Cao *et al*, 2002; Czermin *et al*, 2002; Kuzmichev *et al*, 2002; Muller *et al*, 2002) and the recently discovered EZH1 (Margueron *et al*, 2008; Shen *et al*, 2008). Polycomb group (PcG) proteins are important for regulating the expression of many genes of developmental importance, including the Hox, Pou, Pax, Fox, Sox and Tbx transcriptional regulators (Boyer *et* al, 2006; Lee *et* al, 2006). Thus the loss of the PcG proteins is extremely detrimental to development (Schumacher *et* al, 1996; O'Carroll *et al*, 2001; Voncken *et al*, 2003). The polycomb proteins show a remarkable ability to compact chromatin (Francis *et* al, 2004), therefore it is not surprising that genes bound by these complexes, and tri-methylated at H3K27, are transcriptionally silent. Like mono- and di-methylated H3K9, H3K27me3 is found at silent genes in euchromatin, but H3K27me3 is generally more dispersed throughout the coding region of the gene than methylated H3K9.

Histone methylation is not limited to repressed chromatin. Tri-methylation of H3K4, carried out by hSet1, MLL1 and MLL2 in humans (Milne *et al*, 2002; Wysocka *et al*, 2003; Yokoyama *et al,* 2004) is associated with transcriptionally active genomic regions (Litt *et al*, 2001; Noma *et al*, 2001). H3K4me3 is found at the TSSs of nearly all active genes (Barski *et al*, 2007). Interestingly, the methylation status of H3K4 alters across the coding sequence of a gene, with di-methylated H3K4 being found in the middle of transcribed sequences, and mono-methylated H3K4 being enriched at the 3' end. Tri-methylated H3K4 antagonises the binding of Suv39h1, preventing the methylation of H3K9, and impairs the binding of the NuRD complex (Nishioka *et al*, 2002; Zegerman *et al*, 2002).

The arginine residues of histone proteins can also be methylated. Histone H3 is methylated on arginines 2, 17, 26 (mediated by coactivator associated arginine methyltransferase-1 (CARM1)) and 8 (carried out by protein arginine methyltransferase-5 (PRMT5)), whereas histone H4 is methylated on the arginine residue at position 3 by either PRMT1 or PRMT5 (Figure 1.2). Although CARM1 is involved in nuclear hormone receptor-mediated transcription (An *et al*, 2004),

genome-wide ChIP-chip analysis found that H3R2 methylation is not enriched in active genes over those that are transcriptionally silent (Barski *et al*, 2007). Similarly, the same study found that H4R3 methylation, which stimulates the acetylation of histones H3 and H4 (Huang *et al*, 2005), was not preferentially identified at active gene loci (Barski *et al*, 2007).

Intriguingly, histones have the potential to be modified by both active and repressive marks simultaneously. Studying the histone modifications present at particular regions of mouse ES cell genomes that encode the *Hox* genes and other developmentally important transcription factors, Bernstein *et al* found that three quarters of the loci that have the H3K27me mark also contain sites of H3K4 methylation (Bernstein *et al*, 2006). These marks co-localise in pluripotent ES cells, whereas upon differentiation, most histones have either one mark or the other, as presumably the transcriptional fate of the genes has been decided. Bivalency can also be found in non-ES cells: promoters with a high content of CpG islands in neuroprogenitor cells and mouse embryonic fibroblasts have regions where H3K27 and H3K4 are both methylated, which is resolved to either one of the marks, or no modification upon lineage commitment (Mikkelsen *et al*, 2007). Mass spectrometry has allowed individual *Tetrahymena* nucleosomes to be analysed, showing that H3K4me1 can be found on the same histone tail as H3K27me2, and to a lesser extent, H3K27me1 (Taverna *et al*, 2007), entertaining the possibility that individual histones at key developmental stages are positioned to go down one of a number of potential expression pathways.

Histone H3K4 methylation has been shown to have a unique relationship with histone acetylation. Histones methylated at K4 are subjected to continuous turnover of acetylation, whereas histones methylated on K9 do not share this characteristic (Hazzalin and Mahadevan, 2005). Additionally, nucleosomes in the first intron and exon of the proto-oncogenes *c-fos* and *c-jun* are simultaneously modified by H3ac, H3K4me and H3S10p (see below) (Hazzalin and Mahadevan, 2005). Inhibition of HDACs, which increases the amount of acetylation present at *c-fos* and *c-jun*, but decreases the turnover of this mark, actually inhibited the induction of these genes,

suggesting that at these sequences at least, it is not the presence of acetylated histones that is the important for gene expression, but the dynamic turnover of acetylation (Hazzalin and Mahadevan, 2005).

Histone demethylation is carried out by LSD1 and the jumonji family of proteins (reviewed in Cloos *et al*, 2008). LSD1 demethylates mono- or di-methylated histone H3K4, but cannot demethylate the tri-methylated form of the protein, due to its requirement for a protonated nitrogen molecule as a substrate, which is lost when the histone is tri-methylated (Shi *et al*, 2004). The jumonji domain has no such requirement of its substrate, therefore proteins bearing this domain can remove the tri-methyl mark from histones. For example, the JMJD2 subfamily of proteins (JMJD2A, JMJD2B and JMJD2C) can all demethylate the H3K9me3 mark (Cloos *et al*, 2006; Fodor *et al*, 2006; Klose *et al*, 2006; Whetstine *et al*, 2006), whereas JMJD1A can demethylate the H3K9me1 or H3K9me2 substrates (Yamane *et al*, 2006). JMJD2A and JMJD2B are also capable of demethylating H3K36 (Fodor *et al*, 2006; Klose *et al*, 2006; Whetstine *et al*, 2006), as is Fbxl11 (Tsukada *et al*, 2006). Like LSD1, JARID1C can demethylate lysine 4 of histone H3 (Iwase *et al*, 2007). The methylation of H3K27 is reversed by two proteins, Utx and JMJD3 (Agger *et al*, 2007; De Santa *et al*, 2007; Lan *et al*, 2007; Lee *et al*, 2007a). Finally, arginine residues can also be demethylated, with H3R2 and H4R3 modification being reversed by JMJD6 (Chang *et al*, 2007).

1.1.3.3. Histone Phosporylation

Each of the core histones can be phosphorylated (Figure 1.2). Phosphorylation of H3S10 and H3S28 by Aurora B kinase is associated with mitotic condensation (Nowak and Corces, 2004). The same residues can be phosphorylated by MSK1/2 (mitogen- and stress-activated protein kinase 1 or 2) in an inducible manner during interphase (Clayton and Mahadevan, 2003). Whilst phosphorylation of serine 10 is known to occur at the immediate early genes (Clayton *et al*, 2000; Thomson *et al*, 2001), the location of the serine 28 modification is unknown, and this separation of

targets is evident in the visibly distinct localisation of the modifications by immunofluorescence (Dyson *et al*, 2005). The linker histone H1 is phosphorylated by the cyclin dependent kinases (CDKs), and, like the phosphorylation of H3S10, this modification is associated with chromosome condensation. Although the phosphorylation of H1 during interphase is linked to chromatin relaxation, metaphase chromosomes are hyperphosphorylated on histone H1, thus its exact role is still uncertain (Contreras *et al*, 2003; Roque *et al*, 2008).

Histone phosphorylation is also linked to the DNA double strand breaks (DSBs) through a histone H2A variant called H2AX. The formation of DSBs triggers the phosphorylation of H2AX at S139 in the proximity of the DSB, and this form of the protein is known as _-H2AX. _-H2AX foci accumulate many proteins required for DNA repair, and whilst not essential for repair to occur, cells without _-H2AX cannot cope with DSBs as well as those with it (Thiriet and Hayes, 2005).

1.1.3.4. Histone Ubiquitylation and Sumoylation

Unlike the other modification marks discussed so far, ubiquitylation involves the addition of an entire protein entity to a lysine residue of a histone. Histones H2A and H2B can be mono-ubiquitylated on lysines 119 and 120 respectively (Figure 1.2). On H2A, the modification is catalysed by the murine polycomb proteins Ring1A and Ring1B, members of the PRC1 complex. Following the tri-methylation of H3K27 by the EZH2 component of PRC2, PRC1 binds to this mark and ubiquitylates histone H2A (de Napoles *et* al, 2004; Wang *et al*, 2004a). H2AK119ub is therefore found at transcriptionally silent chromatin. In humans, two enzymes, RNF20/hBRE1 and UbcH6 work together to ubiquitylate H2B (Zhu *et al*, 2005). Although ubiquitylation of histone H2B is generally associated with transcriptionally active regions, knockdown of hBRE1 only affects the expression of a small number of genes (Shema *et al*, 2008). Of those genes that were misregulated upon loss of hBRE1, similar numbers were either up- or down-regulated, suggesting that H2B ubiquitylation can function as either an active or repressive mark (Shema *et al*,

2008). Like other modifications, ubiquitylation has been postulated to function as a point of recruitment for other proteins, although due to its large size (at 76 amino acids and 9kDa, it is almost as large as the histone proteins) it may be capable of physically 'wedging' the chromatin open (Kouzarides, 2007).

Sumoylation involves the modification of lysine residues by the addition of the 11kDa small ubiquitin-like modifier (SUMO) protein. After it was found that H4 can be sumoylated *in vitro* in 293T mammalian cells (Shiio and Eisenman, 2003), it was then shown that each of the core histone proteins can be sumoylated in *Saccharomyces cerevisiae* (*S. cer*) (Nathan *et al*, 2006). Sumoylation of transcription factors is relatively common, and is generally associated with repression (Stielow *et al*, 2008); likewise, sumoylation of histones is believed to decrease transcription levels (Shiio and Eisenman, 2003; Nathan *et al*, 2006).

1.1.3.5. Other Histone Modifications

As well as those modifications outlined above, histones are the substrates of at least three other marks. The linker histone H1 is the main recipient of the poly(ADP-ribose) mark, the addition of which is catalysed by the poly(ADP-ribose) polymerases (PARPs). There has been some indication that poly(ADP-ribosyl)ation of H1 can relax chromatin structure, but this has yet to be proved *in vivo* (Ausió *et al*, 2001). Biotinylation of histones has been linked with gene silencing, cell proliferation, DNA repair and apoptosis (Peters *et al*, 2002; Narang *et al*, 2004; Kothapalli *et al*, 2005). Histones H3 and H4 are subjected to this modification, at lysines 8 and 12 on H4 (Camporeale *et al*, 2004), and lysines 4, 9 and 18 on H3 (Kobza *et al*, 2005). Finally, carbonylation of histones has been shown to occur *in vitro* and *in vivo*, and it appears that the linker histone H1 is the main target of this modification (Wondrak *et al*, 2000).

1.1.4. Histone Binding Proteins

Whilst some histone modifications may directly affect the conformation of chromatin, the link between a modification and its influence on transcription is often not clear until the code has been interpreted by proteins that bind to specific modifications and carry out further alterations to the chromatin. This indirect, effector-mediated method appears to be the main link between histone modifications and gene expression levels. Some of the important histone-binding domains are outlined below, and in Table 1.1.

The bromodomain of the HAT enzyme PCAF was the first histone-binding domain to have its structure resolved, and was found to target acetylated lysine residues (Dhalluin *et al*, 1999). For example, the bromodomain of the *S. cerevisiae* homolog of PCAF, Gcn5p, binds H4K16ac (Owen *et al*, 2000). The largest subunit of the transcription factor complex TFIID, TAFI (TBP-associated factor-1, formerly known as $TAF_{II}250$), encodes a double bromodomain at its C-terminal end, allowing this protein to bind to two acetylated lysine residues simultaneously (Jacobsen *et al*, 2000), directly linking the acetylation of histones to transcription.

The 'Royal family' contains proteins encoding motifs such as the chromo, Tudor and MBT domains, each of which can bind to methylated lysine residues of histones. One of the best characterised histone-binding proteins is heterochromatin protein 1 (HP1), of which there are three isoforms in mammals: HP1_, HP1_ and HP1_, encoded by *Cbx5*, *Cbx1* and *Cbx3* in mouse respectively (Singh *et al*, 1991; Saunders *et al*, 1993). The chromo (chromatin organisation modifier) domain of HP1 was found to bind to H3K9me3 (Bannister *et* al, 2001; Lachner *et al*, 2001), whereas its chromoshadow domain (not a member of the Royal family) facilitates HP1 dimerisation (Cowieson *et al*, 2000). The chromodomain of Polycomb (Pc), highly similar to that of HP1, also binds to methylated lysines, and is selective for H3K27me3, anchoring the PRC1 complex and allowing it to ubiquitylate H2A at the same sites as H3K27me3 (Fischle *et al*, 2003). Another member of the 'Royal' family is the Tudor domain, two of which are found in JMJD2A. This motif allows

Domain Name	Histone Modification	Example	Reference
Bromodomain	Acetyl-lysine	Gcn5p	Owen *et al*, 2000
		TAF1	Jacobsen *et al*, 2000
Chromodomain	Methyl-lysine	HP1	Bannister *et al*, 2001; Lachner *et al*, 2001
		Polycomb	Fischle *et al*, 2003
Tudor	Methyl-lysine	JMJD2A	Huang *et al*, 2006; Lee *et al*, 2008
Plant homeodomain (PHD)	Methyl-lysine	BPTF	Li *et al*, 2006; Wysocka *et al*, 2006
		ING2	Peña *et al*, 2006
Macro domain	Poly(ADP-ribosyl)ated histones	PARP9, PARP14 and PARP15	Karras *et al*, 2005

Table 1.1. Histone binding domains. Table showing histone binding domains found in various chromatin associated proteins, and the histone modifications that they bind to.

the interaction of JMJD2A with two distinct histone modifications, H3K4me3 and H4K20me3 (Huang *et al*, 2006; Lee *et al*, 2008).

H3K4me2 and H3K4me3 are bound by the plant homeodomain (PHD) finger motif of two proteins, bromodomain plant homeodomain transcription factor (BPTF) (Li *et al*, 2006; Wysocka *et* al, 2006) and inhibitor of growth 2 (ING2) (Peña *et al*, 2006). Finally, the macro domain, found in proteins such as the histone variant macroH2A, and three PARPs (PARP9, PARP14 and PARP15) can bind to mono- and poly(ADP-ribose), and it has been postulated that this domain can bind to poly(ADP-ribosyl)ated histones (Karras *et al*, 2005).

These examples show that, although it was originally thought that the modification of histones provided a code by which the transcriptional state of a gene could be ascertained, it now appears that it is instead the proteins that are bound to the modifications that are required to effectively read the genome. The fact that the same histone modifications can be bound by proteins with opposing effects on transcription, that one nucleosome can be modified by marks that are linked to differing expression statuses, and that an enzyme can bind to an active mark and cause repression all suggest that the proteins that bind to histone modifications are the determining factor as to the expression status of a sequence, and not the modifications themselves.

1.1.5. DNA Methylation

In addition to being packaged by histone proteins that are modified, the DNA double helix is subjected to one major modification itself: the addition of a methyl group at the fifth carbon of the cytosine residue in the context of CpG. In vertebrate cells, DNA methylation is widespread: over 70% of CpG dinucleotides are methylated (Meehan, 2003). DNA is methylated *de novo* by two methyltransferases, Dnmt3a and Dnmt3b (Okano *et al*, 1999), and the pattern of methylation is maintained following the synthesis of the new DNA strand by Dnmt1, thus DNA

methylation is a heritable epigenetic mark. The vast majority of DNA methylation is located at the constitutive heterochromatic regions of mammalian genomes: human and mouse centromeres are composed of thousands of satellite repeat sequences, which contain CpG dinucleotides that are heavily methylated. The majority of the remaining CpG dinucleotides are found in the early replicating portions (R-bands) of the genome (Craig and Bickmore, 1994), clustered into domains known as CpG islands which are typically found within gene promoters (Bird *et al*, 1985). Most of the CpG dinucleotides found as part of CpG islands are not methylated, even when the genes they are associated with are transcripitionally silent (Meissner *et al*, 2008).

DNA methylation can prevent efficient gene expression by blocking the association of particular transcription factors with promoters or other binding sites. Alternatively, as with histone modifications, methylated DNA may be bound by proteins that exert a repressive effect on transcription. In mammals, four proteins bearing the methyl binding domain (MBD) (MeCP2 (Meehan *et al*, 1989), MBD1, MBD2 and MBD4 (Hendrich and Bird, 1998)) and three zinc finger proteins (Kaiso (Prokhortchouk *et al*, 2001), ZBTB4 and ZBTB38 (Filion *et al*, 2006)) have been shown to bind methylated DNA. Most of these proteins function as transcriptional repressors, and it is thought that this is through their association with chromatin remodelling complexes or histone modifying enzymes, such as the interactions between MeCP2 and components of the Sin3 HDAC complex (Jones *et al*, 1998; Nan *et al*, 1998), MBD1 and SETDB1 (Sarraf and Stancheva, 2004) and MBD2 and the MeCP1 HDAC complex (Ng *et al*, 1999; Feng and Zhang, 2001). Although it has yet to be proven definitively whether DNA methylation is the cause of, or occurs as a result of histone modification, it is clear that the two modification processes are intrinsically linked.

1.2. Localisation of DNA Within the Nucleus

1.2.1. Organisation of Interphase Chromosomes

At the nucleosomal level, we can see that the organisation of the genome is highly regulated, with histone modifications, nucleosome remodellers and chromatin-binding proteins combining to retain the open or closed nature of euchromatin and heterochromatin, and to allow the expression of genes when required. Looking at the genome as a whole within the nucleus, it is clear that the actual physical localisation of genomic sequences is also regulated. In comparison with their characteristics during metaphase, interphase chromosomes are relatively decondensed, and cannot be distinguished by staining the DNA alone. It might have been expected that the DNA of separate chromosomes is randomly dispersed in interphase; however, visualising individual chromosomes with specific paints revealed that chromosomes from humans and other vertebrates are relatively discrete entities (Figure 1.3A and Figure 1.4). Vertebrate chromosomes were originally thought to be completely separated by an interchromatin space, but recent work has shown that chromosome territories actually intermingle at their peripheries (Branco and Pombo, 2006). Moreover, analysis of the localisation of each chromosome during interphase has shown that they have a conserved radial position, with gene rich chromosomes in general being found at the centre of the nucleus, and gene poor chromosomes lying towards the exterior (Croft *et al,* 1999; Sadoni *et* al, 1999; Boyle *et al*, 2001; Bolzer *et al*, 2005) (Figure 1.3A). There also appears to be a correlation between the radial localisation of a chromosome and its physical size in basepairs, with small chromosomes lying more centrally, and larger chromosomes towards the periphery in non-dividing cells (Sun *et al*, 2000; Bolzer *et al*, 2005). Inhibition of transcription can affect both the amount of intermingling of chromosome territories (Branco and Pombo, 2006) and their size within the nucleus (Croft *et al*, 1999).

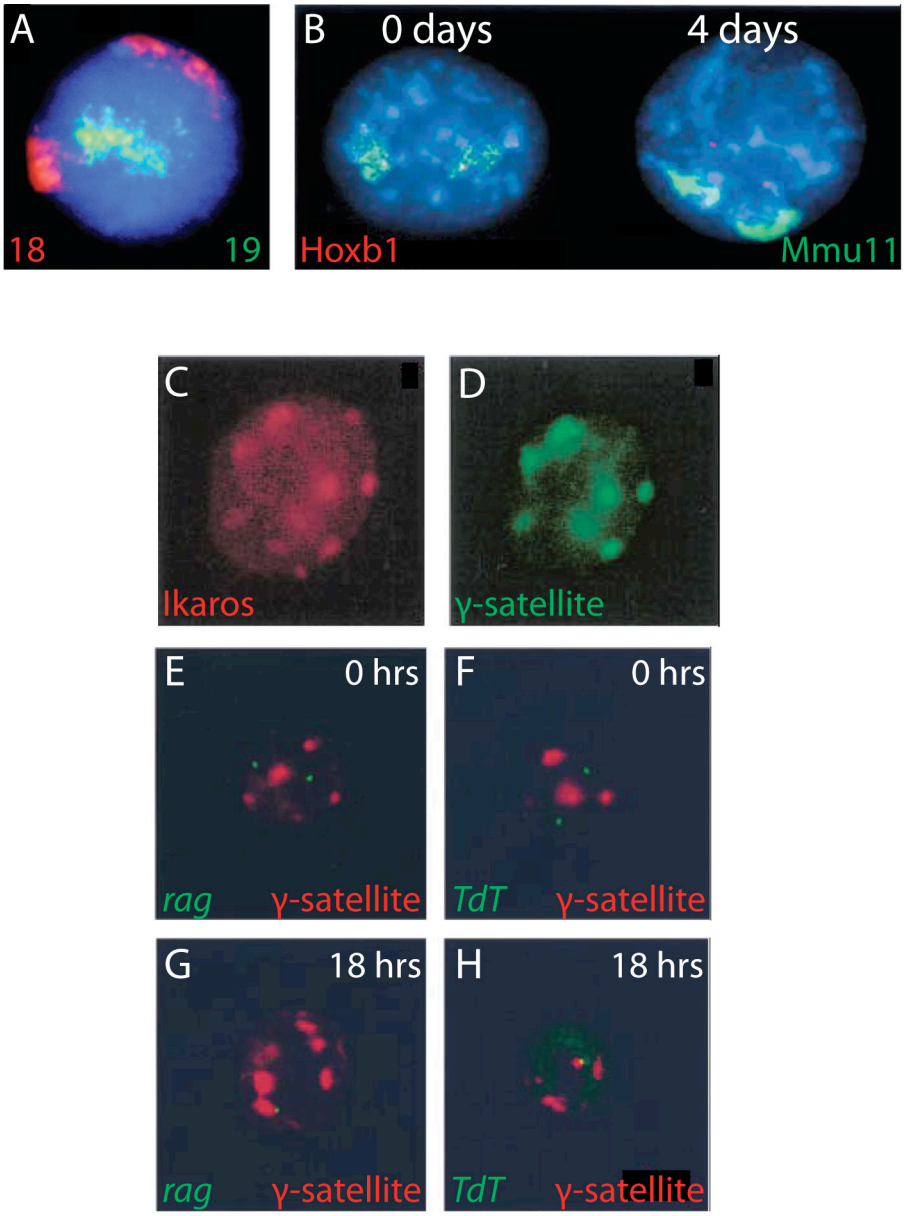


Figure 1.3. Localisation of particular DNA sequences within mammalian nuclei. A) Human chromosomes 18 (red) and 19 (green) are arranged in a non-random manner, with gene-rich chromosome 19 lying more central than gene-poor chromosome 18, shown in primary lymphocytes (Croft *et al*, 1999). B) In mouse ES cells, *Hoxb1* (red) loops out of its chromosome territory (Mmu11; green) upon activation, following 4 days of treatment with retinoic acid (right panel) (Chambeyron and Bickmore, 2004). C) and D) Ikaros protein (C; red) co-localises with mouse γ-satellite (D; green) at heterochromatic foci in mouse B3 cells (Brown *et al*, 1997). E), F), G) and H) *rag* and *TdT* genes are expressed in immature mouse thymocytes, and localise to the nucleoplasm in ~85% of cells tested (E and F). 18 hours of T cell receptor treatment causes the downregulation of *rag* and *TdT*, and these sequences are now found at the heterochromatin in at least of half the cells tested (G and H) (Brown *et al*, 1999).

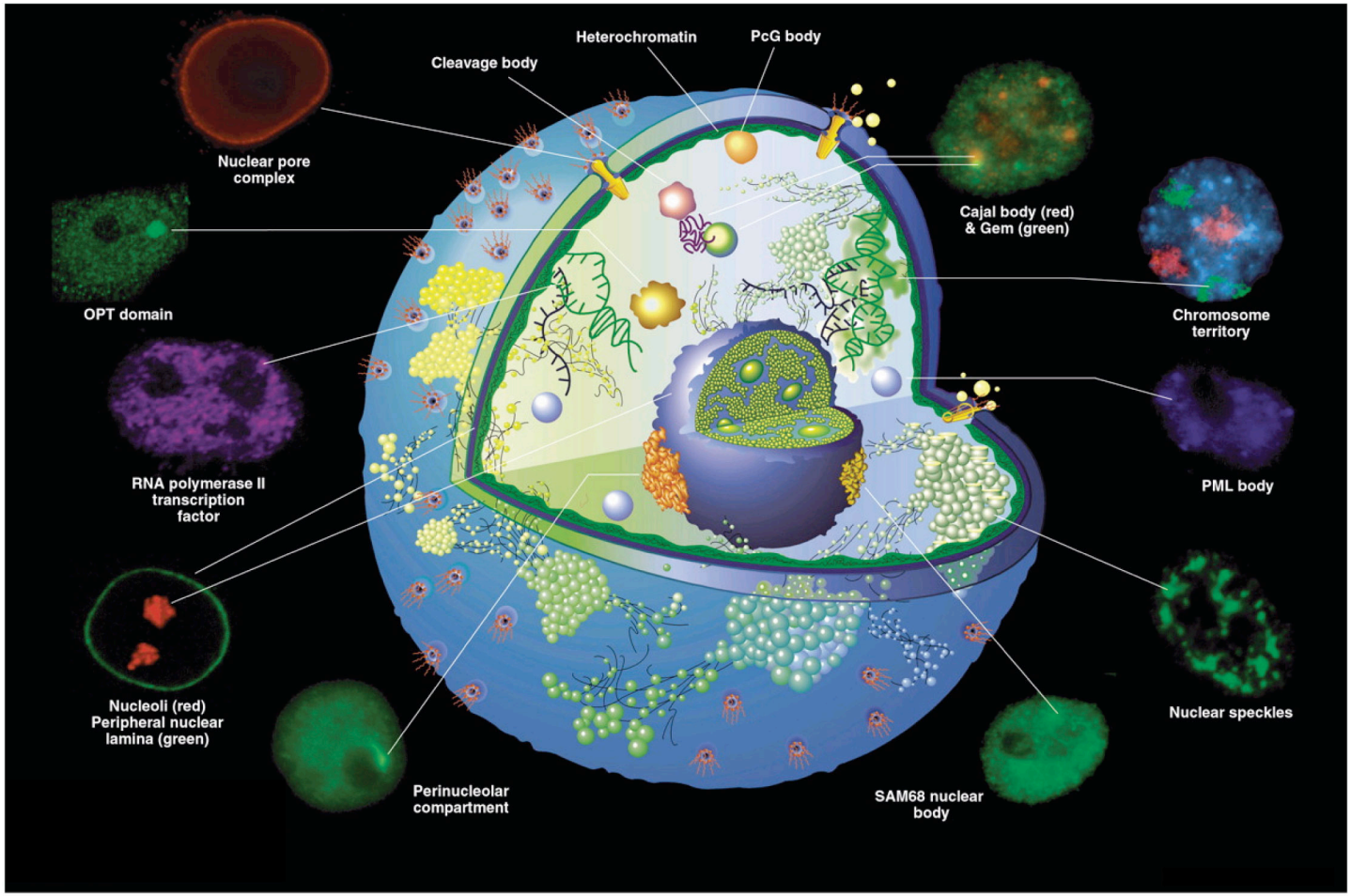


Figure 1.4. Structure of the mammalian cell nucleus. Diagram shows commonly observed nuclear compartments, with examples observed by fluorescence microscopy. PML bodies are seen as distinct stuctures, although these can associate with the Cajal bodies, gems and cleavage bodies. Figure adapted from Spector, 2001.

1.2.2. Relocalisation of Expressed Sequences: Looping Out

The packaging of genes in three-dimensional space with respect to the chromosome territory is altered in some circumstances. Housekeeping genes and those expressed in a tissue-specific manner can be transcribed from the interior of a chromosome territory (Mahy *et al*, 2002), and both active and inactive sequences are found at the periphery of chromosome territories (Kurz *et al*, 1996). Despite this, certain sequences leave their chromosome territory when expressed. For example, the major histocompatibility complex (Volpi *et al*, 2000), the epidermal differentiation complex (Williams *et al*, 2002) and the *Hoxb* locus (Chambeyron and Bickmore, 2004) have all been demonstrated to loop out away from the main body of the chromosome territory when transcribed, leading to the idea that they are relocated to specialised sites containing factors required for their expression (see Figure 1.3B). However, studies of the *Hoxd* locus have shown that looping out is neither necessary nor sufficient for transcription. In ES cells and the mouse tail bud, the *Hoxd* locus loops out, but in the limb bud, where genes from this locus are also activated, decondensation and expression of the *Hoxd* locus is not accompanied by looping out (Morey *et al*, 2007). Similarly, the looping out of a *Hoxb1* transgene inserted into the *Hoxd* locus is not sufficient to induce the transgene's transcription (Morey *et al*, 2008). Positioning outside of the chromosomal territory does not necessarily affect transcriptional status, either. The human _- and _-*globin* genes are located in regions of contrasting gene density (_-*globin* in a gene-rich region, _-*globin* in a gene desert). The _-*globin* locus often lies outside the HSA16 chromosome territory to which it belongs, whereas the _-*globin* locus sits within its chromosome territory (Mahy *et al*, 2002b; Brown *et al*, 2006), but both genes are expressed in the same tissues, at the same rate (Brown *et al*, 2006). Therefore different factors affecting localisation and expression may be at play in different tissues, and gene activation of the same sequences may not always be accompanied by the same hallmarks.

1.2.3. Transcription Factories

Much research in the past few years has looked at where genes that are looping out of their territories, or those that show a high level of expression, localise to. Incorporation of the nucleotide analogue bromouridine (BrdU) into newly synthesised RNA revealed that transcription appears to occur at specific foci containing RNA polymerases and other factors, which are approximately 50nm in diameter (Jackson *et al*, 1993; Iborra *et al*, 1996; Jackson *et al*, 1998). These foci are dispersed evenly throughout the nucleus, except for the nucleolus and the very periphery of the nucleus, agreeing with the finding that transcription can occur from any part of a chromosome territory. Estimates of the numbers of transcription foci are dependent on the cell type, but given that the values range from 1 200 (Brown *et al*, 2008) to 8 000 (Iborra, 2002), and the fact that there could be 65 000 active RNA polymerase II complexes present in the nucleus (Jackson *et al*, 1998), each focus must contain at least 8 active polymerases on average. Therefore these foci were named transcription factories, where multiple transcription units co-localise (Jackson *et al*, 1993). Certain genes, for example the _- and _-*globin* genes in human erythroblasts, show an increased frequency of co-localisation when co-expressed, and it has been suggested that they are recruited to the same transcription factory (Osborne *et al*, 2004). Contradictory work has indicated that active genes are instead juxtaposed at nuclear speckles, and that this is dependent on the status of the chromatin, i.e. the levels of decondensation and transcriptional activity, and its interaction with other proteins (Brown *et al*, 2006; Brown *et al*, 2008). Nuclear speckles are domains, approximately 0.5_m in diameter, rich in poly(A) RNA and splicing factors (Spector, 1993), and have previously been found to be sites where specific pairs of genes and genes on certain chromosomal bands co-localise (Shopland *et al*, 2003; Moen *et al*, 2004). Association with speckles is not required for transcription of these genes, as only a percentage of cells are found to do it, but it may improve the rate of their expression.

Globin genes have also provided much of the evidence that expressed genes come into contact with enhancer sequences. Enhancers are generally highly

conserved regions that can lie tens or hundreds of kilobases from a gene's coding sequence, which enhance transcription by coming into contact with the gene's promoter region. An enhancer element known as the locus control region (LCR) is vital for high-level expression of the _-*globin* genes in mouse erythroid cells (Bender *et al*, 2000). The LCR, situated 50kb away from the murine _-*globin* genes, comes into physical contact with these genes to form an active chromatin hub (Bender *et al*, 2000; Tolhuis *et al*, 2002). Formation of the hub is tissue-specific (Palstra *et al*, 2003) and dependent upon certain transcription factors (Drissen *et al*, 2004; Vakoc *et al*, 2005; Splinter *et al*, 2006). Further analysis of the LCR showed that this region can also contact loci tens of megabases (Mb) away, and that the LCR in liver cells (where the _-*globin* genes are expressed) associates with regions containing transcriptionally active genes, whereas the LCR in brain cells (where the _-*globin* locus is transcriptionally inactive) preferentially associates with silent genomic regions (Simonis *et al*, 2006). A ubiquitously expressed gene on a different chromosome, *Rad23a*, showed the same interactions in both liver and brain tissue, and in both cases interacted with transcriptionally active loci (Simonis *et al*, 2006). Although it appeared that the transcription status of the LCR was the determining factor in the 'choice' of interactions that the LCR makes, transcription itself is not required for the maintenance of these interactions (Palstra *et al*, 2008). This suggests that the associations that a particular locus makes are instead dependent on the factors bound to the interacting sequences. Certain loci appear to have a preferential set of interactions with other regions of the genome, thus the folding of a chromosome does not seem to be a completely random process, but one which preserves the interactions that specific loci make.

1.2.4. Sites of Transcriptional Repression Within the Nucleus

The constitutive heterochromatin of mammalian cells, composed of the tandemly repeated satellite repeats, forms highly condensed structures. These AT-rich structures are visible in mouse nuclei upon staining with 4',6-diamidino-2-phenylindole (DAPI). The DAPI bright spots are also rich in H3K9me3 and certain

isoforms of HP1. HP1_ is found to be highly enriched at the DAPI bright spots, whereas HP1_ localises to both the constitutive heterochromatin and the euchromatin, and HP1_ is primarily found at euchromatic regions (Horsley *et al*, 1996; Minc *et al*, 1999; Minc *et al*, 2000; Minc *et al*, 2001). Genes that are physically associated with the heterochromatic bright spots are often transcriptionally inactivated. The Ikaros protein, which is expressed in haematopoietic cells, is found to localise to the pericentromeric heterochromatin in some lymphocyte cell lines (Figure 1.3C-D; Brown *et al*, 1997; Brown *et al*, 1999) and to recruit genes that are transcriptionally repressed, such as *TdT* and *Rag* (Brown *et al*, 1999) to these nuclear sites (Cobb *et al*, 2000) (Figure 1.3E-H). The human _-*globin* gene is also observed at the centromeric heterochromatin when silenced, but loses this association in cells where it is transcriptionally active (Brown *et al*, 2001). Interestingly, the _-*globin* locus in the same cells is never found at the centromeric heterochromatin (Brown *et al*, 2001), which most likely reflects its chromosomal positioning within a cluster of ubiquitiously expressed housekeeping genes, whose silencing would be detrimental to the cell.

Heterochromatin is also found at the very edge of the nucleus underlying the nuclear membrane, and attached to the matrix surrounding the nucleolus (Figure 1.4; reviewed in Comings, 1980). The interior surface of the nuclear periphery is made up of a family of intermediate filament proteins called the lamins, which form a complex called the nuclear lamina. The inner nuclear membrane (INM) is also decorated with specific proteins. Lamins and INM proteins can bind DNA and histones *in vitro* (Luderus *et al*, 1992; Taniura *et al*, 1995), and this area of the nucleus has therefore been postulated to have a role in chromatin organisation and gene regulation. DNA associated with lamin B1 at the nuclear periphery is enriched in gene deserts, is predominantly transcriptionally silent, deficient for methylated H3K4 and RNA polymerase II, and enriched in methylated H3K27 and H3K9 (Guelen *et al*, 2008). Proteins at the nuclear periphery are also known to interact with repressive proteins and complexes such as HP1, N-CoR and HDACs (Kourmouli *et al*, 2000; Somech *et* al, 2005; Holaska and Wilson, 2007). Consistent with this, some genes localise to the edge of the nucleus when repressed. During B-lymphocyte

development, the immunoglobulin (Ig) loci are subject to rearrangement and transcription to form and express different Ig proteins. Murine *IgH* and *Ig_* loci are situated at the nuclear periphery as a default in most cell types; however, in pro-B lymphocyte cells, these loci become positioned towards the nuclear interior prior to their transcription, suggesting that their retention at the periphery keeps them in a repressed state (Kosak *et al*, 2002). Two separate studies, analysing the human *CFTR* locus (Zink *et al*, 2004) and the mouse *Mash1* gene (Williams *et al*, 2006), found that when repressed, these genes localise to the nuclear periphery, but that in an expressed state, they move towards the interior of the nucleus (Zink *et al*, 2004). However, as with looping out upon gene activation, localisation to the nuclear periphery does not ensure transcriptional silencing of a DNA sequence. The *interferon-_* (*IFN-_*) gene, which is upregulated upon the differentiation of CD4 T cells, remains at the nuclear periphery in the majority of cells tested regardless of its transcriptional activity (Hewitt *et al*, 2004).

The link between localisation to the nuclear periphery and transcriptional repression has recently been tested more rigorously by three independent studies that used the same tactic of artificially tethering genes to the nuclear periphery in mammalian cells and analysing their subsequent expression (Finlan *et al*, 2008; Kumaran and Spector, 2008; Reddy *et al*, 2008). Finlan *et al* showed that transcription of an integrated reporter gene was decreased when tethered to LAP2_, an INM protein. In the regions of chromosomes 4 and 11 surrounding the reporter, some endogenous genes were subjected to downregulation, but many genes were also unaffected by their relocalisation (Finlan *et al*, 2008). Kumaran and Spector found that reporter gene expression was still possible upon recruitment to the periphery in the majority of cells tested, although compared to control cells with an untethered reporter, the number of alleles being actively transcribed was decreased (Kumaran and Spector, 2008). Reddy *et al* reported that transcription of their reporter was significantly repressed upon association with the nuclear periphery (Reddy *et al*, 2008). It therefore appears that localisation to the nuclear periphery is not a generalised method of silencing, and the lamina does not provide an entirely

transcription-free environment, although certain genes are repressed when localised there.

1.2.5. Other Domains of the Nucleus: PML Bodies

The nucleus is home to many other protein structures that are important for different aspects of gene expression and the cell cycle (Figure 1.4). One such structure is the promyelocytic leukaemia nuclear body, or PML-NB. PML (also known as tripartite motif protein 19, or Trim19) was first identified as being associated with acute promyelocytic leukaemia (APL) (de Thé *et al*, 1991). A specific translocation between human chromosomes 15 and 17 fuses PML with the *retinoic acid receptor _* (*RAR_*) locus, preventing promyelocyte differentiation and causing APL (de Thé *et al*, 1991). PML is a member of the tripartite motif protein family, containing an N-terminal RING, B-boxes and coiled-coil (RBCC) domain or tripartite motif (Figure 1.5). In the middle of the protein lies a nuclear localisation signal (NLS), and the C-terminus harbours an exonuclease domain (Salomoni *et al*, 2008). There are at least seven different splice variants of the *PML* gene, each of which creates a different PML isoform through variation in the length of the C-terminal end of the protein (Salomoni *et al*, 2008). Each isoform contains the RBCC domain and NLS, but only PML-I contains the exonuclease motif (Figure 1.5). Despite their differing characteristics and functions, these variants will be hereafter collectively referred to as PML. In the nucleus, PML is found as part of the large, matrix-associated domains named PML-NBs, nuclear domain 10 (ND10), Kremer (Kr) bodies or PML oncogenic domains (PODs) (Figure 1.4). The average number of PML-NBs per cell is 10 (hence the original name, ND10 (Ascoli and Maul, 1991)), but there can be between 5 and 30 PML bodies in the nucleus, which appears to be cell type- and cell cycle-dependent (Dellaire and Bazett-Jones, 2004). Mutation of each portion of the tripartite motif of PML can affect NB formation. Altering amino acids of the RING domain increases the size, and reduces the number, of PML-NBs in mouse cells (Boddy *et al*, 1997), whereas mutation of the B-box domains prevents PML-NB formation (Borden *et al*, 1996).

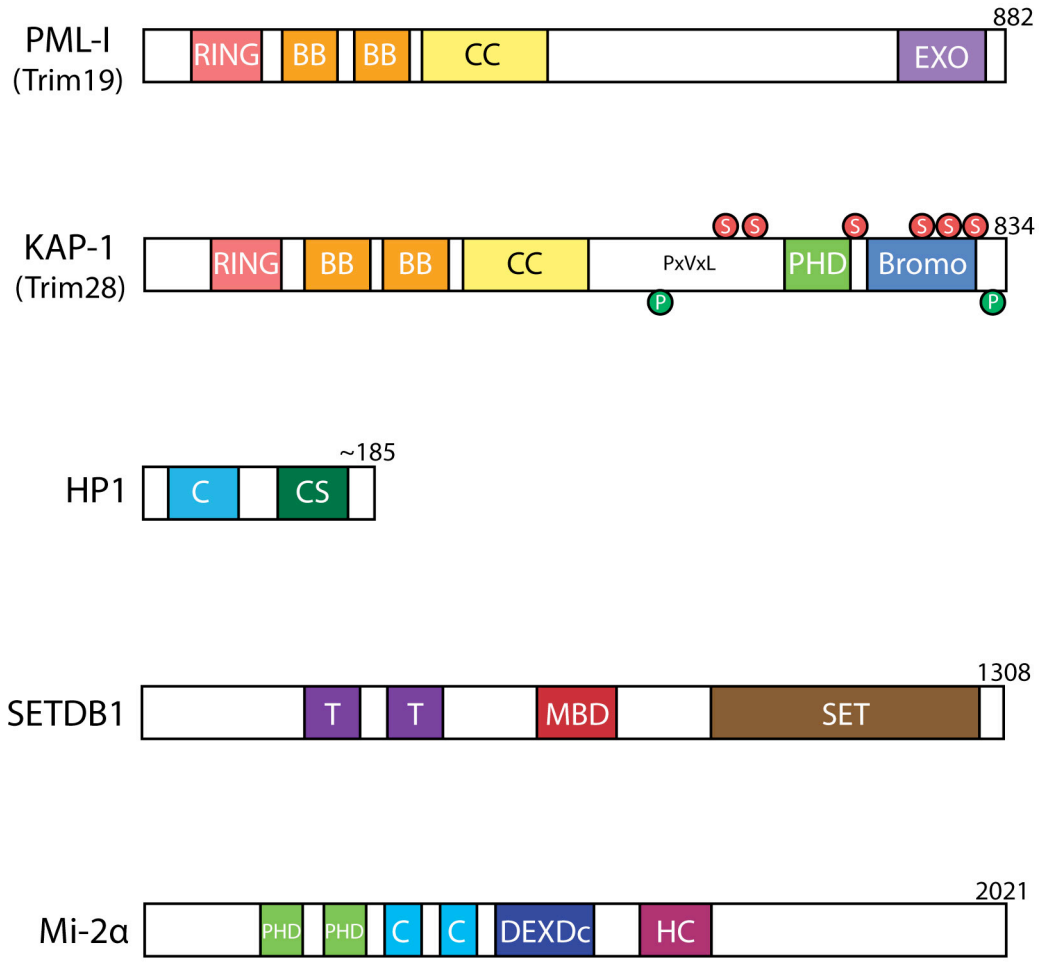


Figure 1.5. Structure of the tripartite motif proteins PML and KAP-1, and the KRAB zinc finger protein-associated factors. Diagram showing the domains found in PML, KAP-1, HP1, SETDB1 and Mi-2α. Sites of sumoylation (red circles) and phosphorylation (green circles) on KAP-1 are shown. Diagrams are not to scale. RING: really interesting new gene; BB: B-boxes; CC: coiled-coil; EXO: exonuclease domain; PxVxL: HP1-binding motif; PHD: plant homeodomain; Bromo: bromodomain; C: chromodomain; CS: chromoshadow domain; T: Tudor; MBD: methyl-binding domain; SET: Suppressor of Variegation 3-9, Enhancer of zeste and trithorax; DEXDc: DEAD-like helicases superfamily; HC: helicase superfamily C-terminal domain.

PML-NBs are highly dynamic in structure and composition. They are typically 0.3-1.0_m in diameter (Ascoli and Maul, 1991; Dellaire and Bazett-Jones, 2004) and either spherical or ring-shaped (Boisvert *et al*, 2000; Lallemand-Breitenbach *et al*, 2001; Eskiw *et al*, 2003). PML-NBs do not appear to have any nucleic acid at the core of their structure (Boisvert *et al*, 2000), but associate with sites of transcription (Kießlich *et al*, 2002), heterochromatin (Seeler *et al*, 1998; Luciani *et al*, 2006), centromeres (Everett *et al*, 1999), telomeres (Yeager *et al*, 1999) and damaged DNA (Xu *et al*, 2003; Boe *et al*, 2006; Dellaire *et al*, 2006). Through its interaction with SATB1 (special AT-rich sequence binding protein 1), PML has been shown to organise the higher-order chromatin structure of the major histocompatibilty complex (MHC) class I locus, and can thus regulate the expression of the genes encoded at this locus (Kumar *et al*, 2007). As well as SATB1 and other proteins involved in gene regulation, PML-NBs contain proteins that belong to various cellular pathways, including DNA repair, antiviral response, apoptosis and tumour suppression (Dellaire and Bazett-Jones, 2004). Because of the diverse mixture of proteins located there, PML bodies have been suggested to be either storage depots for nuclear proteins (Maul *et al*, 2000; Negorev *et al*, 2001), sites allowing the post-translational modification of proteins (D'Orazi *et al*, 2002), or hubs where particular events occur, for example DNA replication or repair (Dellaire and Bazett-Jones, 2004). However, knockout of PML is not lethal: $PML^{-/-}$ mice have a normal appearance, but are deficient in their apoptotic response to DNA damage (Wang *et al*, 2004b).

Whatever their exact function, it is clear that PML-NBs are intrinsically linked to post-translational modification by SUMO. PML was one of the first proteins identified to be modified in this manner (Boddy *et al*, 1996), and while sumoylation of PML was found to be non-essential for the formation of PML-NBs, it is required for the accumulation of other nuclear factors such as Sp100 and Daxx at PML-NBs (Zhong *et al*, 2000). PML encodes a SUMO-interacting motif (SIM), which can bind to other SUMO-modified PML moieties, and it is believed that this is how orderly PML-NB structures are formed (Shen *et al*, 2006). Other proteins often present at PML-NBs, for example Daxx, are also modified by SUMO. Like PML, Daxx too encodes a SIM. While the localisation of Daxx to PML-NBs is independent of its

sumoylation, a functional SIM is required for its recruitment to these sites (Lin *et al*, 2006).

The formation of PML-NBs requires the presence of wild-type PML (Ishov *et al*, 1999). In APL patients, PML-NBs are either disrupted or dispersed (Dyck *et al*, 1994). Treatment of APL patients with either retinoic acid (RA) or arsenic oxide (As_2O_3) can cause the destruction of the PML-RAR_ fusion protein, leading to the differentiation and apoptosis of cells affected by APL and remission of the disease (Zhu *et al*, 2002). Upon treatment of cells with As_2O_3, PML is targeted for polysumoylation. Poly-sumoylated PML is then poly-ubiquitinated by the ubiquitin E3 ligase RNF4, and the protein (either native PML in tissue culture cells or PML-RAR_ in APL patients) is destroyed by the proteasome, again sending some APL patients into remission (Lallemand-Breitenbach *et* al, 2008; Tatham *et al*, 2008).

As well as their association with APL, PML bodies can be linked to cancer through their role in the alternative lengthening of telomeres (ALT) pathway. This occurs in cancerous cells that retain their telomere length by homologous recombination (HR) to increase their proliferative capacity. Telomeres in these cells are recruited to the PML-NBs, forming ALT-associated PML-NBs (APBs) (Yeager *et al*, 1999). APB formation and telomeric HR depends on the structural maintenance of chromosomes 5/6 (SMC5/6) complex, which harbours SUMO ligase activity through its MMS21 subunit (Potts and Yu, 2007). MMS21-mediated sumoylation of the shelterin complex, which normally protects the telomeres from homologous recombination, is required for the recruitment of telomeres to APBs, again highlighting the association of PML-NBs and sumoylation.

1.3. KAP-1 and the KRAB Zinc Finger Family of Transcriptional Repressors

1.3.1. KAP-1 and Transcriptional Repression

Mammals encode 37 other tripartite motif proteins in addition to PML. One of the most studied of these is KAP-1 (KRAB-associated protein-1), also known as Trim28, Tif1_ (transcription intermediary factor 1_) or KRIP-1 (KRAB-interacting protein-1). As well as the RBCC domain, KAP-1 encodes a PHD and a bromodomain at its C-terminus (Figure 1.5). Three related proteins are known with this structure: Tif1_ (aka Trim24), Tif1_ (Trim33) and Tif1_ (Trim66). When tethered to a reporter gene, all four of these proteins can cause transcriptional repression (Friedman *et al*, 1996; Kim *et al*, 1996; Le Douarin *et al*, 1996; Moosmann *et al*, 1996; Venturini *et al*, 1999; Khetchoumian *et al*, 2004); however the sequence identity between these proteins is not high, at between 30 and 50% (Khetchoumian *et al*, 2004). Tif1_ is tightly associated with chromatin, is expressed early in mouse development, where it has been implicated in the control of the first wave of transcriptional activation (Remboutsika *et al*, 1999; Torres-Padilla and Zernika-Goetz, 2006), and is also expressed in adult tissues (Le Douarin *et al*, 1995). Tif1_ (KAP-1) is expressed ubiquitously throughout development (Cammas *et al*, 2000) and in adult tissues (Kim *et al*, 1996). Knockout of KAP-1 results in the termination of development prior to gastrulation (Cammas *et al*, 2000), demonstrating the essential nature of this protein. Tif1_ is also expressed in both embryonic and adult tissues (Venturini *et al*, 1999; Yan *et al*, 2004), whereas Tif1_'s expression is limited to the testis (Khetchoumian *et al*, 2004). Tif1-_, -_ and -_ can all homodimerise; Tif1_ and Tif1_ can also heterodimerise *in vitro* (Peng *et al*, 2002; Germain-Desprez *et al*, 2003).

KAP-1 was independently identified by three groups as the co-repressor of a family of proteins called the Krüppel-associated box (KRAB) zinc finger proteins (Friedman *et al*, 1996; Kim *et al*, 1996; Moosmann *et al*, 1996). These proteins encode a KRAB domain at their N-terminus, and most also have multiple zinc finger

motifs at their C-terminal end. The KRAB zinc finger proteins are not capable of repressing transcription themselves, but are believed to bind to specific gene targets through their zinc fingers and exert a repressive effect through the recruitment of KAP-1 by the KRAB domain. This interaction occurs through the RBCC domain of KAP-1, which is also responsible for KAP-1 oligomerisation (Figure 1.6; Peng *et al*, 2000). Different portions of this N-terminal region have been implicated in the KRAB-KAP-1 interaction; it remains unclear whether it is the entire RBCC domain (Peng *et al*, 2000), the RING domain and B-boxes (Friedman *et al*, 1996) or the B-boxes and coiled coil domains (Moosmann *et al*, 1996; Agata *et al*, 1999) that are vital. The RING finger co-ordinates two zinc ions, producing a 'cross-brace' motif unique to this domain (Meroni and Diez-Roux, 2005). Many E3 ubiquitin ligases, which catalyse the transfer of a ubiquitin moiety from an E2 ligase to a substrate protein (or commonly to itself) contain a RING domain, and the RING sequence has been shown to be critical in this reaction (Joazeiro and Weissman, 2000). Thus any proteins that contain a RING domain may have E3 ligase activity. The ~40 amino acid B-boxes encoded in KAP-1 are known as B-box 1 and B-box 2. Like the RING domain, they are thought to bind zinc (Borden *et al*, 1995), but no specific function has been assigned to these structures. The coiled-coil region stretches over 126 amino acids, and is involved in KAP-1 homo-interaction (Reymond *et al*, 2001).

KAP-1 causes transcriptional silencing through the proteins it interacts with. One of the first studies to identify KAP-1 did so as a result of its ability to bind to HP1 in a yeast two-hybrid assay (Le Douarin *et al*, 1996). Further experimentation showed that this interaction occurred through the central portion of KAP-1, from amino acids 483-510, and that a KRAB domain, the DNA it is bound to, KAP-1 and HP1 could form a quaternary complex *in vitro* (Figure 1.5 and Figure 1.6; Ryan *et al*, 1999). It was later revealed that this section of the protein harbours an HP1-binding motif with the sequence PRVSL, adhering to the consensus of PxVxL (Smothers and Henikoff, 2000) which binds to the HP1 chromoshadow domain (Lechner *et al*, 2000) (Figure 1.6). Loss of the interaction between HP1 and KAP-1 leads to a reduction in the KRAB-KAP-1 mediated repression of both transfected and chromatinised reporter genes (Ryan *et al*, 1999; Sripathy *et al*, 2006). In undifferentiated cell lines, KAP-1

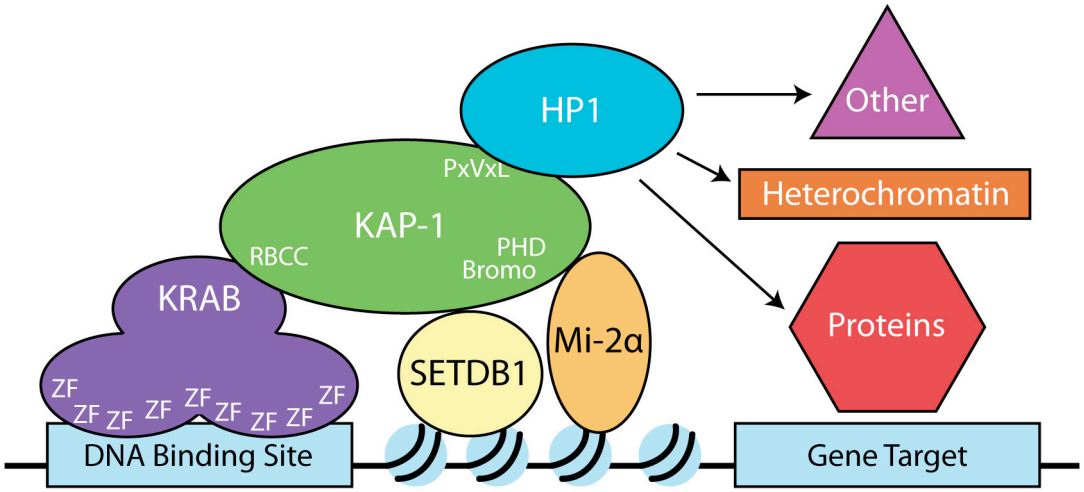


Figure 1.6. Mechanism of KRAB zinc finger protein-mediated transcriptional repression. KRAB zinc finger proteins are believed to bind to specific gene targets through their zinc finger motifs, and recruit the KAP-1 co-repressor via their KRAB domain. KAP-1 in turn can bind to HP1 through its PxVxL motif, and the chromatin modifying proteins SETDB1 and Mi-2α via its PHD and bromodomains. Together, these proteins can create a silent chromatin environment, stifling expression of the KRAB zinc finger protein's gene targets.

is diffuse throughout the nucleus, and excluded from the nucleolus. Upon differentiation of F9 embryonic carcinoma (EC) or ES cells, however, the protein often relocalises to the centromeric heterochromatin, showing a similar distribution pattern to HP1_ (Cammas *et al*, 2002; Cammas *et al*, 2004). By mutating the PxVxL consensus sequence, it was shown that the binding of KAP-1 to HP1 is required for the protein's relocalisation, and that this interaction is actually essential for cell differentiation (Cammas *et al*, 2002; Cammas *et al*, 2004).

At the carboxy terminus of KAP-1 are a PHD finger and a bromodomain. As well as being able to bind to histones, PHD and bromodomains can bind to non-histone chromatin proteins, for example the PHD and bromodomain of NoRC, which interacts with HDAC1, SNF2 and two DNA methyltransferases (Zhou and Grummt, 2005). The bromodomain of Tif1_ interacts directly with nucleosomes *in vitro* (Remboutsika *et al*, 2002); however, this property has not been shown for KAP-1 to date. Examination of the role of the C-terminal region of KAP-1 revealed that the PHD and bromodomain can function as transcriptional repressors when tethered to DNA, and that the two domains in concert bind to the chromatin modifying proteins Mi-2_ (Schultz *et al*, 2001) and SETDB1 (Schultz *et al*, 2002) (Figure 1.5 and Figure 1.6). Mi-2_ is a component of the NuRD HDAC and chromatin remodelling complex, which is most commonly linked to transcriptional repression (Denslow and Wade, 2007). The association of the KRAB-KAP-1 machinery with this complex explains why treatment of cells with trichostatin A, a HDAC inhibitor, reduces KAP-1's repressive ability (Nielsen *et al*, 1999; Schultz *et al*, 2001). SETDB1 encodes an MBD and a bifurcated SET domain, and is known to specifically methylate histone H3K9 (Schultz *et al*, 2002).

1.3.2. Evolution of KRAB Zinc Finger Proteins

When first identified, the 75 amino acid KRAB domain was estimated to be present in approximately one third of all human proteins encoding Krüppel-type zinc fingers, making it one of the largest known protein families (Bellefroid *et al*, 1991).

Despite being present in ~400 human proteins (Huntley *et al*, 2006) and a similar number in mouse, the KRAB domain is restricted to the vertebrate tetrapods, with a decreased number (six) in *Xenopus laevis* (the most studied of these being Xfin (Ruiz i Altaba *et al*, 1987; Bellefroid *et al*, 1991)) and 41 KRAB proteins predicted in chicken (http://smart.embl-heidelberg.de/). This suggests that the KRAB zinc finger proteins have undergone a relatively recent, and extremely large, evolutionary expansion within the mammals. As might be expected from a gene family that has recently expanded, half of the human and mouse KRAB zinc finger proteins are found in large clusters: 11 major clusters in the human genome are home to 200 KRAB zinc finger protein loci (Shannon *et al*, 2003). The human ZNF91 subfamily is one of the subgroups of the KRAB zinc finger proteins. (Human KRAB zinc finger proteins are denoted by the three-letter acronym ZNF; murine KRAB zinc finger proteins are denoted by Zfp). These genes can be found in two major clusters, with 49 genes on HSA19 and 27 on either side of the centromere of HSA7 (Hamilton *et al*, 2006). These genes initially began duplicating approximately 40 million years ago, have seeded new clusters on other chromosomes due to genomic rearrangements, and have continued to expand, giving rise to a number of lineage-specific genes (Hamilton *et al*, 2006). The other half of the KRAB zinc finger proteins do not belong to gene clusters, and are thought to be less functionally redundant than their clustered counterparts. Despite their abundance in the mammalian genome, however, remarkably little is known about the role of the KRAB zinc finger proteins.

Although many KRAB zinc finger proteins are conserved between species, there has been gain, loss and modification of KRAB zinc finger protein genes since the generation of their clusters. For example, the same KRAB zinc finger protein cluster was analysed in human and mice on HSA19q13.2 and MMU7, encoding 21 and 10 genes respectively (Shannon *et al*, 2003) (Figure 1.7). Three genes were identified which were conserved in a 1:1 manner. One mouse gene had given rise to ten human genes, and one human gene had six mouse counterparts, suggesting selective pressure has lead to the retention and modification of some sequences and the loss of others (Shannon *et al*, 2003). Many of the KRAB zinc finger protein genes at these

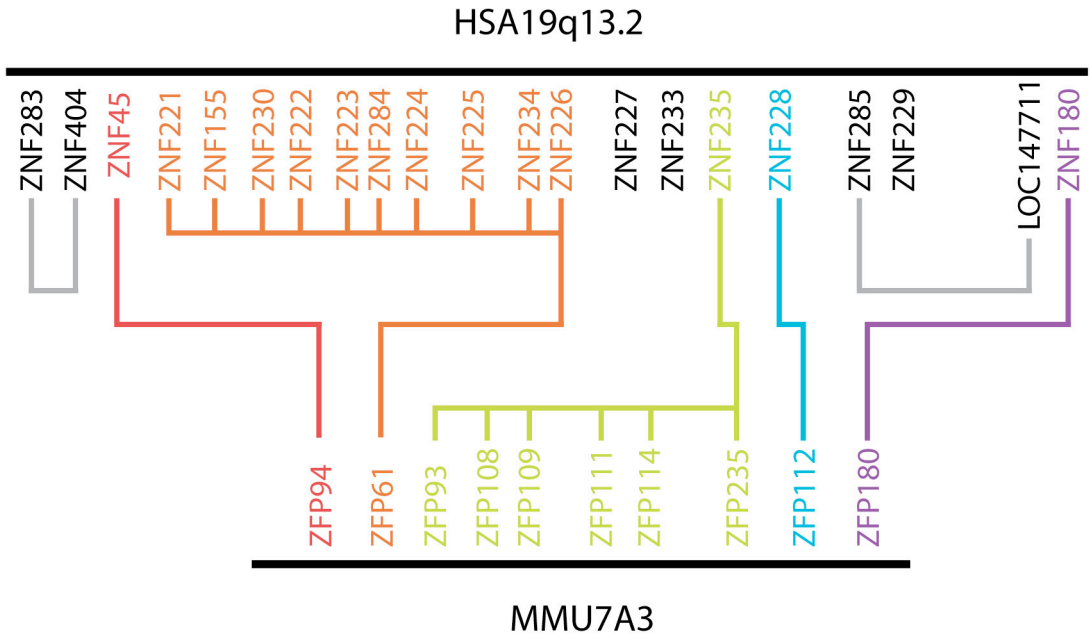


Figure 1.7. A comparison of two KRAB zinc finger protein gene clusters. Diagram of homologous KRAB zinc finger protein clusters at HSA19q13.2 (a 700kb cluster encoding 21 genes) and MMU7A3 (a 300kb cluster encoding 10 genes). The conservation and expansion of gene ancestors is shown by the connecting lines between genes from the two species. The relationships between the genes shown in black have not been as well defined as those in colour, but high levels of homology between two pairs of genes are indicated by the grey lines. Adapted from Shannon *et al*, 2003.

large loci have evolved to a greater degree in the sequences encoding their zinc finger domain compared to the KRAB domain. This will have affected the DNA sequences to which they can bind, but not the mechanism by which they cause transcriptional repression. Thus, although there are a great number of KRAB zinc finger proteins, the degree of redundancy between them may be limited, as they have separate gene targets. Therefore, as it appears that different species have a slightly different collection of KRAB zinc finger protein genes, they might be of importance in the emergence of species-specific gene expression, and the evolution of each species as it is today.

Whilst high numbers of KRAB zinc finger proteins are restricted to mammals, other lineages appear to have undergone large-scale expansion of particular zinc finger proteins. *Drosophila*, like humans, encode many C2H2 zinc finger proteins, believed to be between 234 (Venter *et al*, 2001) and 352 (Rubin *et al*, 2000). 89 of these proteins contain a ~70 amino acid N-terminal domain, termed the zinc finger-associated C4DM domain, or ZAD (Lander *et al*, 2001; Chung *et al*, 2002; Lespinet *et al*, 2002). Analysis of other insect, non-insect invertebrate and vertebrate genomes showed that these ZAD-associated zinc finger proteins are restricted to five insect species: *Drosophila melanogaster*, *Anopheles gambiae*, *Aedes aegyptii*, *Apis mellifera* (honey bee) and *Bombyx mori* (silk moth). Three mutant ZAD alleles have been identified in *Drosophila*: *deformed wing/zeste-white5* (*dwg/zw5*) (Fahmy and Fahmy, 1959), *grauzone* (Schupbach and Wieschaus, 1989) and *Serendipity* _ (Payre *et al*, 1990), and studies on the encoding proteins have suggested that they are involved in the regulation of transcription. Like the KRAB zinc finger proteins, ZAD-associated zinc finger proteins are not evenly dispersed throughout the genome, with 41 genes being present in clusters. The majority of the ZAD proteins are expressed, inferring that they have retained some function (Chung *et al*, 2002).

1.3.3. Structure of KRAB Zinc Finger Proteins

The KRAB domain was initially shown to act as a potent transcriptional repressor when targeted to reporter gene promoters (Witzgall *et al*, 1993; Margolin *et al*, 1994; Pengue *et al*, 1994). This repression was still observed when the KRAB domain was bound 700bp upstream of the promoter sequence (Pengue *et al*, 1994). The KRAB motif is always situated at the amino terminus of a protein. There are four sub-types of KRAB domain, depending on which of four subdomains are present (Figure 1.8). All KRAB domains contain a KRAB-A box. This can be paired with a KRAB-B (Bellefroid *et al*, 1991) or KRAB-b box (Mark *et al*, 1999), or with the more recently discovered KRAB-C subunit (Looman *et al*, 2004). Alternatively, some KRAB zinc finger proteins only encode the KRAB-A domain. Each of the KRAB subdomains is encoded by a separate exon, and alternative splicing may lead to the removal of the KRAB-B, –b or -C boxes from the protein, increasing the number of distinct KRAB zinc finger proteins that can be expressed (Bellefroid *et al*, 1993). For most KRAB zinc finger proteins, the entire zinc finger domain is encoded by one exon, which, given that the domain may contain 30 zinc finger motifs spanning over one kilobase, is highly unusual, and may suggest that this part of the KRAB zinc finger protein genes evolved through retroposition.

The KRAB-A and KRAB-B motifs both reduce transcription levels, and it has been suggested that the KRAB-B subdomain has a stronger repression activity than the KRAB-A domain (Vissing *et al*, 1995), despite the KRAB-A box being the only indispensable portion of the KRAB motif (Margolin *et al*, 1994). The presence of the highly divergent KRAB-b subdomain does not appear to enhance the overall repression activity of the KRAB domain (Abrink *et al*, 2001). While the exact functions of the KRAB-B and KRAB-b subdomains are unknown, the mechanism of transcription repression has been determined for the KRAB-A motif: it is this domain that binds to KAP-1 (Friedman *et al*, 1996; Kim *et al*, 1996; Moosmann *et al*, 1996). The KRAB-C subdomain enhances the interaction between the KRAB domain and KAP-1, but this apparently does not increase the KRAB zinc finger protein's ability to repress transcription (Abrink *et al*, 2001; Looman *et al*, 2004). In addition, KRAB

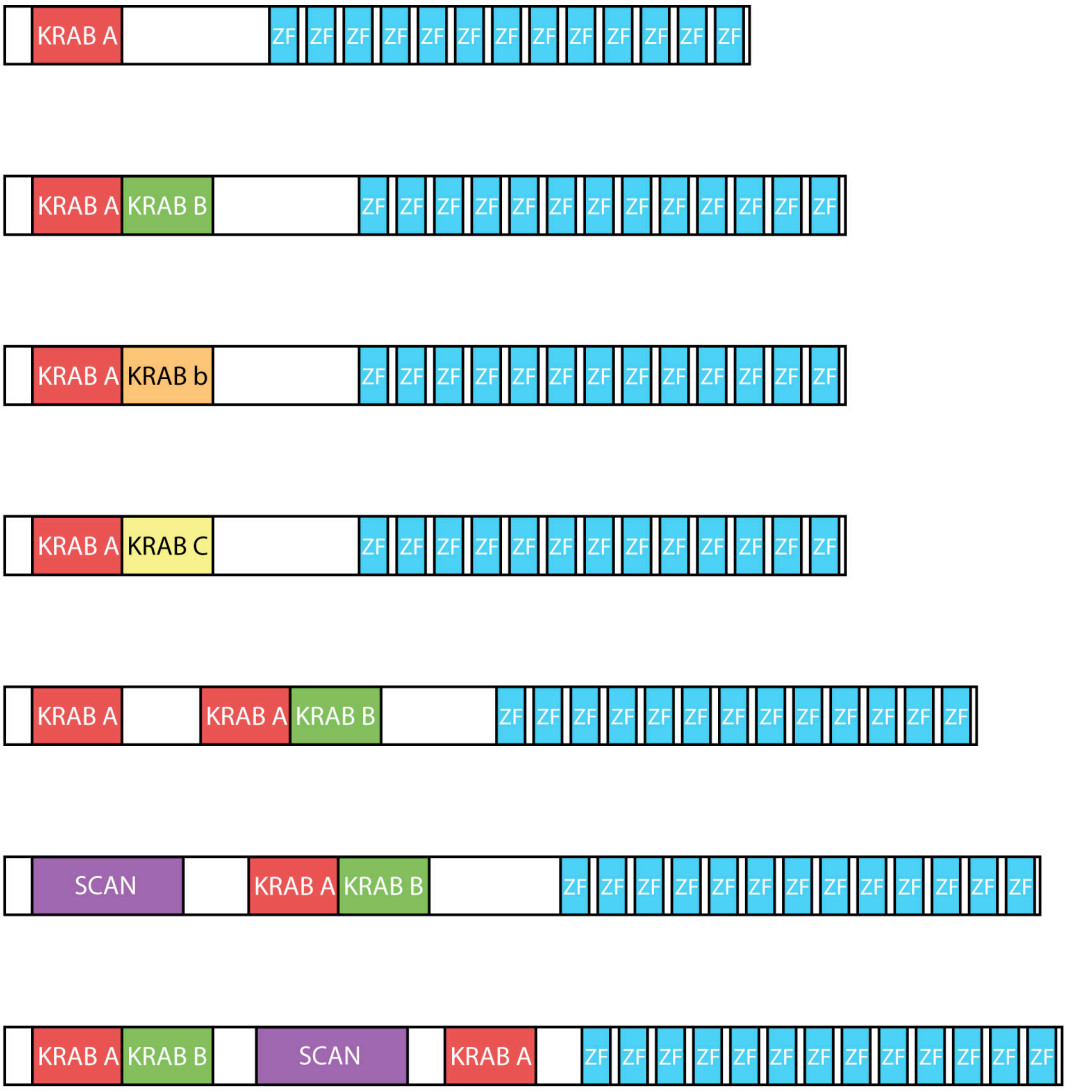


Figure 1.8. Diagram of typical KRAB zinc finger protein structures. Each protein encodes at least one KRAB-A subdomain. KRAB zinc finger protein genes also encode either a KRAB-B, KRAB-b or KRAB-C subdomain, which may be removed by alternative splicing to leave the KRAB-A box alone. KRAB zinc finger proteins also occasionally contain a SCAN domain at their amino-terminal end. Adapted from Urrutia, 2003.

zinc finger proteins occasionally contain a leucine-rich (LeR), or SCAN (SRE-ZBP, CTfin51, AW-1 and Number 18 cDNA) domain (Figure 1.8). This ~90 amino acid structure is normally located beside the KRAB domain, and allows homo- and hetero-dimerisation with other SCAN domain-containing proteins (Honer *et al*, 2001).

1.3.4. Mechanism of KRAB Zinc Finger Protein-Mediated Repression

The mechanism by which KRAB zinc finger proteins cause transcriptional repression is outlined in Figure 1.6. The KRAB zinc finger protein is thought to bind to its target through an unspecified number of zinc finger motifs. This may not require the full complement of zinc fingers, given the large numbers of motifs in some proteins, and the fact that each zinc finger can bind to 3 nucleotides. The KRAB domain interacts with the KAP-1 co-repressor, which acts as a scaffold protein: it can recruit the NuRD HDAC and chromatin remodelling complex, the SETDB1 HMTase, or HP1 proteins. These proteins can cause the heterochromatinisation of the gene target (Sripathy *et al*, 2006), and can bring in factors such as DNA methyltransferases and histone modifying enzymes which could maintain a heterochromatic state at the target locus, keeping it stably repressed.

It is unknown whether KRAB zinc finger proteins cause heterochromatinisation of their targets *in situ* in the cell, or whether they are relocated to a repressive environment within the nucleus to cause gene silencing, for example to the heterochromatin or the nuclear periphery. In support of the idea that KAP-1-mediated repression may occur at sites of constitutive heterochromatin, it has been reported that two KRAB zinc finger proteins, KRAZ1 and KRAZ2, co-localise with HP1_ at the pericentromeric heterochromatin, and that this interaction is dependent on the binding of the KRAB zinc finger proteins to KAP-1 (Matsuda *et al*, 2001). KAP-1 also appears to concentrate at the constitutive heterochromatin during differentiation (Cammas *et al*, 2002), and the tethering of a KRAB-KAP-1 complex

to a reporter plasmid can cause its relocalisation to the DAPI bright spots (Ayyanathan *et al*, 2003).

Evidence against the localisation of KRAB zinc finger protein targets to heterochromatin comes from the fact that SETDB1 predominantly localises to the euchromatic regions of the nucleus, and colocalises with HP1 at non-pericentromeric regions of chromatin (Schultz *et al*, 2002). It was also demonstrated that Mi-2_ can be recruited to a euchromatic gene through its interaction with KAP-1, and thus it was proposed that KAP-1 initiates the formation of a heterochromatin scaffold around the gene, rather than taking the gene to a pre-formed repressive complex (Schultz *et al*, 2001). KAP-1 localises to both heterochromatic and euchromatic regions of the nucleus, which does not aid in deciphering how and where it functions in KRAB-mediated silencing.

1.3.5. Post-Translational Modification of KAP-1 and KRAB Zinc Finger Proteins

KAP-1 is subjected to various post-translational modifications, which are believed to alter its function (Figure 1.5). Firstly, as has been found for many transcription factors, KAP-1 can be modified by SUMO. Out of 42 lysines present, 6 are targets for sumoylation, all of which lie in and around the PHD and bromodomains in the protein's carboxy terminus (Lee *et al*, 2007b; Mascle *et al*, 2007) (Figure 1.5). Three of these sites, K554, K779 and K804, are the major recipient of the mark, and the three other lysines, at positions 575, 676 and 750, are not modified *in vivo* to the same degree (Lee *et al*, 2007b; Mascle *et al*, 2007). Prevention of sumoylation leads to a loss of KAP-1-mediated transcriptional repression at the KRAB zinc finger protein ZBRK1 target gene *p21*, which is accompanied by an increase in lysine acetylation and a decrease in H3K9 methylation at this sequence (Lee *et al*, 2007b). Oligomerisation was also shown to be important for KAP-1 sumoylation, as the removal of the RBCC domain, which co-ordinates KAP-1-KAP-1 interactions, reduces the sumoylation of the protein, but

replacing this domain with another capable of self-interaction restores KAP-1 modification (Mascle *et al*, 2007). Following these two reports, it was shown that KAP-1's PHD-finger (which shows strong similarity to a RING domain) and bromodomain are required for KAP-1 sumoylation, and in fact bind to the SUMO E2 ligase Ubc9, directing modification to KAP-1's bromodomain, thus acting as the protein's E3 ligase (Ivanov *et al*, 2007, Zeng *et al*, 2008). Sumoylation of KAP-1 was also shown to be critical for KRAB-mediated repression, as SETDB1 and Mi-2_ bind to the sumoylated form of the protein *in vivo* (Ivanov *et al*, 2007).

One paper that has surveyed the sumoylation of human proteins in tissue culture cells has suggested that the KRAB zinc finger proteins may also be sumoylated *in vivo*. Vertegaal *et al* isolated proteins that had been modified by exogenously expressed His_6-tagged-SUMO-1 or -SUMO-2, and identified the sumoylated proteins by mass spectrometry (Vertegaal *et al*, 2006). Of the 25 SUMO-1 targets, 14 encoded zinc finger motifs, and two of these proteins were in fact KRAB zinc finger proteins (retinoblastoma (RB)-associated KRAB repressor (RBAK) and ZNF317). KAP-1 was not identified in this screen. Given its close association with the SUMO ligase KAP-1, it is tempting to speculate that KRAB zinc finger protein sumoylation may also play a part in KRAB-mediated transcriptional repression.

KAP-1 is also phosphorylated at two points within its structure, on serines 473 and 824 (Figure 1.5). Phosphorylation of S824 occurs after UV-irradiation of cells or induction of DSBs, and is dependent on the ataxia-telangiectasia mutated (ATM) protein kinase (White *et al*, 2006; Ziv *et al*, 2006). The chromatin of cells treated to induce DSBs is decondensed, an effect that is exacerbated when wild-type KAP-1 is swapped for a phosphorylation mimetic S824D isoform. This suggested that KAP-1 is involved in global chromatin condensation, which is over-ridden by S824 phosphorylation to allow the relaxation of the chromatin (Ziv *et al*, 2006). KAP-1 phosphorylation at this amino acid, and the protein's sumoylation, have been shown to have a negative effect on each other, whereby sumoylation leads to a decrease in phosphorylation of KAP-1, and *vice versa*, and active removal of SUMO leads to an increase in phosphorylated KAP-1 levels (Li *et al*, 2007). If KAP-1-mediated

chromatin decondensation plays an important part in the DSB repair pathway, this phosphorylation mark would allow the conversion of KAP-1 involved in KRAB-mediated gene silencing to the DSB repair isoform without having to remove the protein from the cell.

KAP-1 that is phosphorylated on S473 seems to be present during S and M phases of the cell cycle (Chang *et al*, 2008). This modification is carried out by member of the protein kinase C (PKC) pathway. A S473A mutation, which cannot be phosphorylated, generates a KAP-1 protein that can interact more robustly with HP1, and so it has been postulated that this modification may regulate gene repression carried out by KAP-1 by altering its binding relationship with the heterochromatin proteins (Chang *et al*, 2008).

1.3.6. Regulation of KRAB Zinc Finger Protein Genes by KAP-1 and HP1

Two recent studies have suggested that some of the major targets of KRAB zinc finger proteins and KAP-1 are the genes encoding the KRAB zinc finger proteins themselves. Using ChIP on chip, O'Geen *et al* looked for the location of KAP-1 binding within the human genome (O'Geen *et al*, 2007). By high resolution mapping, the group found that KAP-1 binding sites are enriched for zinc finger transcription factors, and in particular, the KRAB zinc finger protein genes: 212 of the 355 genes were in the list of KAP-1 targets in Ntera2 cells. The length of DNA sequence bound by KAP-1 moieties was 820bp on average, and these sites were concentrated at the 3' ends of the KRAB zinc finger protein genes. Interestingly, this report correlated with another paper, which used the DamID method to map where in the genome HP1_ localises to. HP1_ is also found at the KRAB zinc finger protein genes, and, like KAP-1, was enriched towards the 3' end of the coding sequence, and absent from the 5' end (Vogel *et al*, 2006). HP1_ was found in 20 large domains of up to 4Mb, thousands of times the size of the KAP-1 domains. The majority of the HP1 binding sites coincided with the KRAB zinc finger protein gene clusters. Using the mechanism outlined above, it would appear that KAP-1 binds to the zinc finger gene

clusters, recruits HP1 and the HMTase SETDB1, causing the methylation of H3K9. This would allow the binding of more HP1 moieties, which could in turn facilitate the spreading of H3K9me3 and HP1 to regions outwith that of KAP-1 binding.

These papers suggested two reasons for why the KRAB zinc finger protein genes are themselves subjected to KAP-1/HP1-mediated heterochromatinisation. Firstly, Vogel *et al* argued that HP1 binding and the subsequent condensed chromatin structure protected the KRAB zinc finger protein genes from HR-driven gene removal, therefore these gene clusters were allowed to flourish where they might have otherwise been removed (Vogel *et al*, 2006). O'Geen *et al* analysed the expression of the KRAB zinc finger protein genes, and saw that those bound by KAP-1 were reduced in their expression when compared to all other genes (O'Geen *et al*, 2007). Additionally, they reported that an unspecified number of KRAB zinc finger protein genes were expressed at a high level in Ntera2 cells. Using this evidence, they put forward the idea that the KRAB zinc finger protein genes may be regulated by the KRAB-KAP-1 mechanism. They argued that certain KRAB zinc finger proteins are expressed in particular cell lines, tissues or developmental stages to control the transcription of the remaining KRAB genes. Because little is known about the targets of the KRAB zinc finger proteins, we are unable at this point to ascertain whether proteins with similar functions are released from KRAB-mediated transcriptional repression and expressed simultaneously, although it is easy to imagine that this might be the case.

1.3.7. Gene Targets of KRAB Zinc Finger Proteins

Bona fide gene targets of KRAB zinc finger proteins have not been forthcoming. Most of the work that has shown that KRAB zinc finger proteins and KAP-1 function as transcriptional repressors has measured the expression of reporter genes rather than an endogenous target sequence. The first genuine targets of KRAB zinc finger proteins to be identified were the major urinary protein (*Mup*) genes (Johnson *et al*, 1995); *Cyp2d9*, a cytochrome P450 gene (Sueyoshi *et al*, 1999); and the gene

encoding sex-limited protein (*Slp*) (Georgatsou *et al*, 1993; Tullis *et al*, 2003), which are all typically only expressed in male mice. The activity that represses these genes in female mice (*Regulator of sex-limitation*, or *Rsl*) localises to mouse chromosome 13 (Jiang *et al*, 1996), and was mapped to a 23-gene cluster encoding KRAB zinc finger proteins (Krebs *et al*, 2003). Two of the genes from this cluster, renamed *Rsl1* and *Rsl2*, are required for repression of the male-specific genes in females. The restoration of function of these proteins in mice mutant for both *Rsl1* and *Rsl2* showed that *Rsl1* represses *Slp* and *Cyp2d9*, whereas *Rsl2* inhibits the expression of the *Mup* genes (Krebs *et al*, 2003).

More recently, it was discovered that the KRAB zinc finger proteins are linked to the restriction of retroviral replication in ES and EC cells. Restriction is thought to happen only in these cells in order to prevent germ-line insertional mutagenesis. KAP-1 binds to the primer binding site of an integrated murine leukaemia virus, inhibiting transcription of the viral genes, and thus stopping viral replication (Wolf and Goff, 2007, Wolf *et al*, 2008a). Mutation of the HP1 binding site found in KAP-1 prevented restriction, showing that HP1 is essential for the transcriptional repression of the integrated virus (Wolf *et al*, 2008b). Finally, purification and analysis of the repressive complex showed that a single KRAB zinc finger protein, Zfp809, was responsible for recognising the primer binding site and recruiting KAP-1 (Wolf and Goff, 2009). It is therefore possible that other KRAB zinc finger proteins have evolved to bind to and silence other retroviral sequences. If so, this function may help to explain the expansive evolution of KRAB zinc finger proteins in mammals, as these animals are targeted by a wide range of retroviruses.

1.3.8. Protein-Protein Interactions of KRAB Zinc Finger Proteins

As well as binding to KAP-1, KRAB zinc finger proteins are able to interact with other proteins, which has given some clue as to the range of roles that they play within the cell. Protein-protein interactions can be mediated through the KRAB, linker and zinc finger regions, indicating that each domain may be multi-functional.

The murine protein NRIF (neurotrophin receptor interacting factor), which encodes two KRAB domains, a SCAN domain and five zinc fingers, binds to the neurotrophin receptor $p75^{NTR}$ through its linker region (Casademunt *et al*, 1999), and to another $p75^{NTR}$-interacting protein, TRAF6 (tumour necrosis factor (TNF) receptor-associated factor 6) through its most amino-terminal KRAB domain (Gentry *et al*, 2004). $p75^{NTR}$ promotes programmed cell death in embryonic tissues; loss of NRIF in these cells leads to a significant reduction in cell death (Casademunt *et al*, 1999). Moreover, the interaction between NRIF and TRAF6 which is required for the nuclear localisation of NRIF, is also necessary for the TRAF6-mediated activation of c-Jun NH_2-terminal kinase (JNK). Another protein that forms interactions through its linker region is RBAK. RB is known to repress the expression of genes normally under the control of the E2F transcription factor family (Hamel *et al*, 1992; Hiebert *et al*, 1992). By uncovering its interaction with RB, Skapek *et al* were able to show that RBAK also represses the activation of an E2F-dependent promoter, although the repression activities of RB and RBAK were not in this case co-operative (Skapek *et al*, 2000). Thus it is possible to discover genes that are repressed (or activated) by KRAB zinc finger proteins by analysis of their protein binding partners, rather than trying to directly identify the sequences that they bind to.

KRAB zinc finger proteins can also bind to other factors through their zinc finger domains. KRIM-1 (KRAB box interacting with Myc-1) interacts with c-myc, a transcriptional transactivator, through one of its zinc fingers, which negatively regulates c-myc function (Hennemann *et al*, 2003). Another protein, Zkscan17, binds to the SET domain-containing histone methyltransferase NSD1 through a unique zinc finger derivative called the C2HR motif (Nielsen *et al*, 2004). Zkscan17 can also interact with the jumonji domain-containing protein JARID2, which is critical for cardiovascular development (Mysliwiec *et al*, 2007), indicating that KRAB zinc finger proteins are capable of binding to chromatin remodelling machinery directly. These two reports produced conflicting results as to whether Zkscan17 assists or antagonises transcription. In the case of NSD1, Zkscan17 repressed the reporter construct it was targeted to, and mutation of the C2HR motif suppressed this repression (Nielsen *et al*, 2004). Conflictingly, the paper reporting on Zkscan17's

interaction with JARID2 found that the KRAB zinc finger protein increased transcriptional activity of a reporter construct, which was reduced upon over expression of the jumonji protein (Mysliwiec *et al*, 2007).

Another protein that was shown to form protein-protein interactions through its C-terminus is ZBRK1 (zinc finger and BRCA1-interacting protein with a KRAB domain-1). This protein was identified as a binding partner of BRCA1 (breast cancer-1, early onset), a tumour suppressor that has a role in DNA damage repair. Binding to BRCA1 occurs through the last 4 of the protein's 8 zinc fingers and its extended C-terminal region, which is present in only three other KRAB zinc finger proteins (ZNF577, ZNF613, and FLJ12644) (Zheng *et al*, 2000; Tan *et al*, 2004). ZBRK1 can bind to DNA sequences with the consensus GGGxxxCAGxxxTTT (Zheng *et al*, 2000). One of BRCA1's targets, *GADD45a*, encodes a sequence similar to the ZBRK1 consensus in its third intron, and it has been shown that ZBRK1 can repress the transcription of *GADD45a* through BRCA1, independent of its KRAB domain (Zheng *et al*, 2000). Furthermore, ZBRK1's C-terminal domain can also mediate its homo-tetramerisation, an ability that no other KRAB zinc finger protein has been reported to possess to date. The deletion of 9 amino acids from the carboxy terminus abolishes homo-oligimerisation, and almost completely removes ZBRK1's co-repressor ability, thus self-interaction seems to be of great importance to this protein (Tan *et al*, 2004). Therefore, despite the dearth of verified gene targets of KRAB zinc finger protein, elements of their functions are slowly coming to light through identification of their non-KAP-1 binding partners.

1.4. KRAB Zinc Finger Protein 647 (Zfp647)

1.4.1. Identification of Zfp647

In 2001, the Bickmore lab published a gene trap screen undertaken to identify proteins that localise to specific regions within the mouse nucleus (Sutherland *et al*, 2001). For this screen, a _-*galactosidase-neomycin phosphotransferase* (_-*geo*) vector, designed to integrate into the introns of expressed genes, was transfected into EC and ES cells. This lead to the creation of cell lines that expressed fusions of the N-terminal portion of an endogenous protein attached to the _-*geo* reporter. Selection for G418 resistance ensured that only those clones expressing a correctly integrated and folded fusion protein would survive. X-gal staining of the gene-trapped cells identified those clones where the fusion protein was located within the nucleus, and immunostaining against _-galactosidase was used to determine the exact nuclear localisation of the gene-trapped protein. Cells in which the protein was observed at a specific nuclear compartment, such as the nucleolus, splicing speckles or heterochromatin, were selected, and the trapped genes were identified by 5' RACE. As part of this screen, seven KRAB zinc finger proteins were trapped and identified. One of these proteins, zinc finger protein 647 (Zfp647), was analysed further by Dr Heidi Sutherland and Dr Stephanie Briers.

The *Zfp647* gene is located on mouse chromosome 15, arm q, band D3. Its 535 amino acid sequence encodes both KRAB-A and KRAB-B subdomains, and 13 zinc finger motifs separated from the KRAB domain by a 95 amino acid linker (Figure 1.9A). The gene encoding the homologous protein in humans, *ZNF250*, is conserved in its syntenic location on human chromosome 8, at HSA8q24.3 (Figure 1.9B-C). The KRAB domain and linker region of the human and mouse proteins share 89% and 86% identity at the amino acid level, whereas the zinc fingers are more highly conserved, with 98% identity. This suggests that Zfp647 and ZNF250 may share their DNA binding targets, and thus could regulate the transcription of the same genes. Unlike some of the KRAB zinc finger protein genes, *Zfp647* and *ZNF250* do not belong to a gene cluster, and do not appear to have significant homology to any

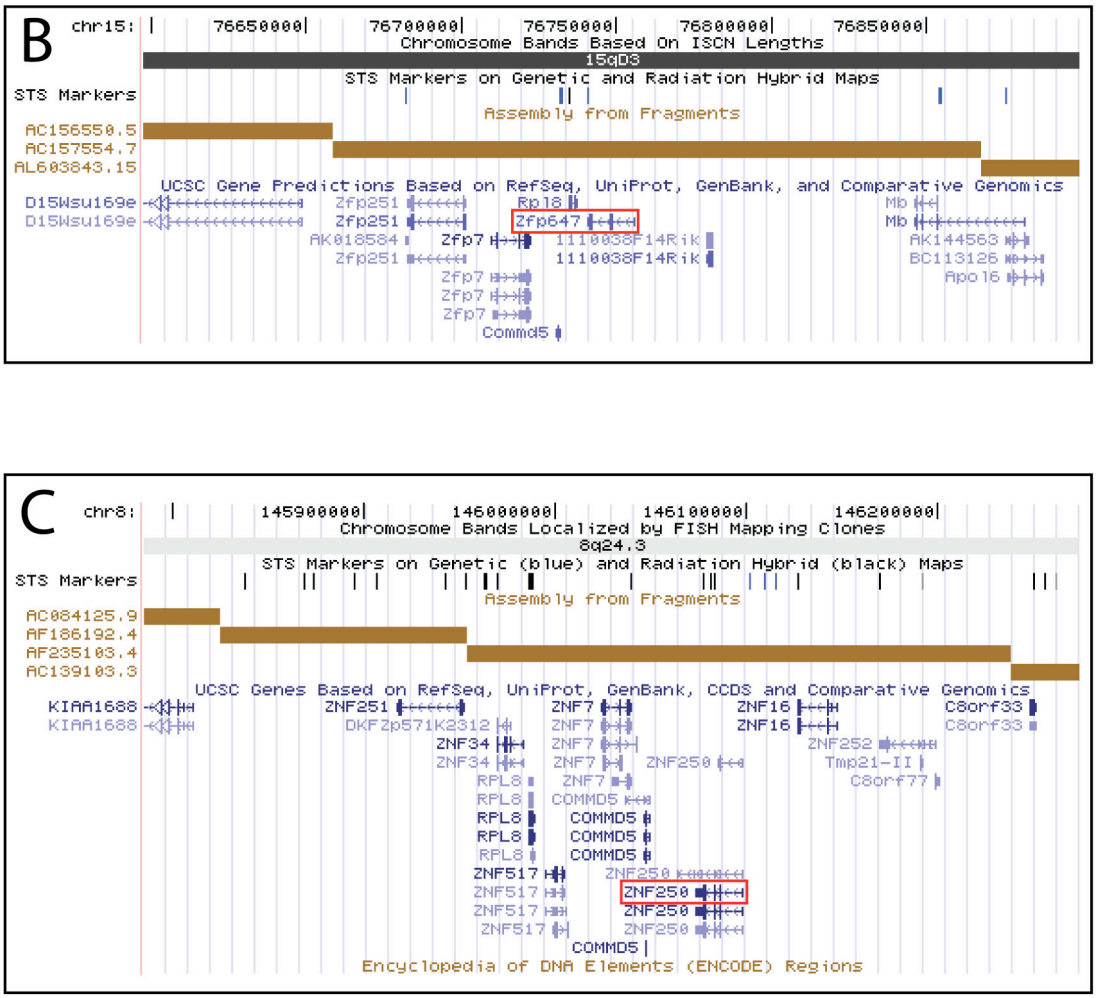


Figure 1.9. Structure and genomic location of Zfp647. A) Protein structure of Zfp647, where KRAB represents the Krüppel-associated box, and ZF represents zinc finger motifs. Numbers indicate the first and last amino acids of the KRAB and zinc finger domains. B) Location of the *Zfp647* gene on mouse chromosome 15, outlined in red. C) Location of the *ZNF250* gene on human chromosome 8,

other KRAB zinc finger proteins, suggesting that their functions are non-redundant. A few KRAB zinc finger protein genes are found in the vicinity of *Zfp647* and *ZNF250* (Figure 1.9B-C); however, the proteins encoded by these genes are less similar to Zfp647 and ZNF250 than other KRAB zinc finger proteins encoded in the mouse and human genomes. In the gene-trapped ES cells, instead of inserting into an intron, the _-*geo* cassette inserted into the 5' end of the final exon encoding Zfp647's zinc fingers (Figure 1.10A). The Zfp647-_-geo fusion protein was still expressed in frame, however, and therefore must use a cryptic splice site. Since the gene-trap insertion is upstream of the zinc fingers, the trapped Zfp647 has probably lost its DNA binding capacity.

1.4.2. Localisation of Zfp647 and KAP-1 in ES Cells

Immunofluorescence of the gene-trapped ES cells with an anti-_-gal antibody showed that in undifferentiated cells, Zfp647 had a diffuse pattern of nuclear staining, appearing in fine speckles throughout the nucleoplasm, but excluded from the nucleoli and the pericentric heterochromatin (Figure 1.10A, -RA). Upon differentiation of the cells with RA, the protein re-localised to the heterochromatic bright spots, detected by DAPI staining, in almost half of the cells analysed. Moreover, this localisation overlapped with that of KAP-1 suggesting that, like the KRAZ proteins, Zfp647 may be recruited to the heterochromatin (Figure 1.10A, +RA; Briers *et al*, 2009). Because the localisation of gene-trapped Zfp647 may have been affected by the loss of its DNA binding domain, a GFP-Zfp647 fusion protein was also generated (Figure 1.10B), and transfected into NIH/3T3 cells. Like the gene-trapped Zfp647 protein, GFP-Zfp647 localised to the pericentric heterochromatin in 20% of the cells analysed (Figure 1.10B, upper fluorescence panel), and also in some cases to small non-heterochromatic foci (Figure 1.10B, lower fluorescence panel).

However, the sub-cellular localisation of the GFP-fusion protein may differ from that of the endogenous protein. Thus two antibodies were raised by Dr Stephanie

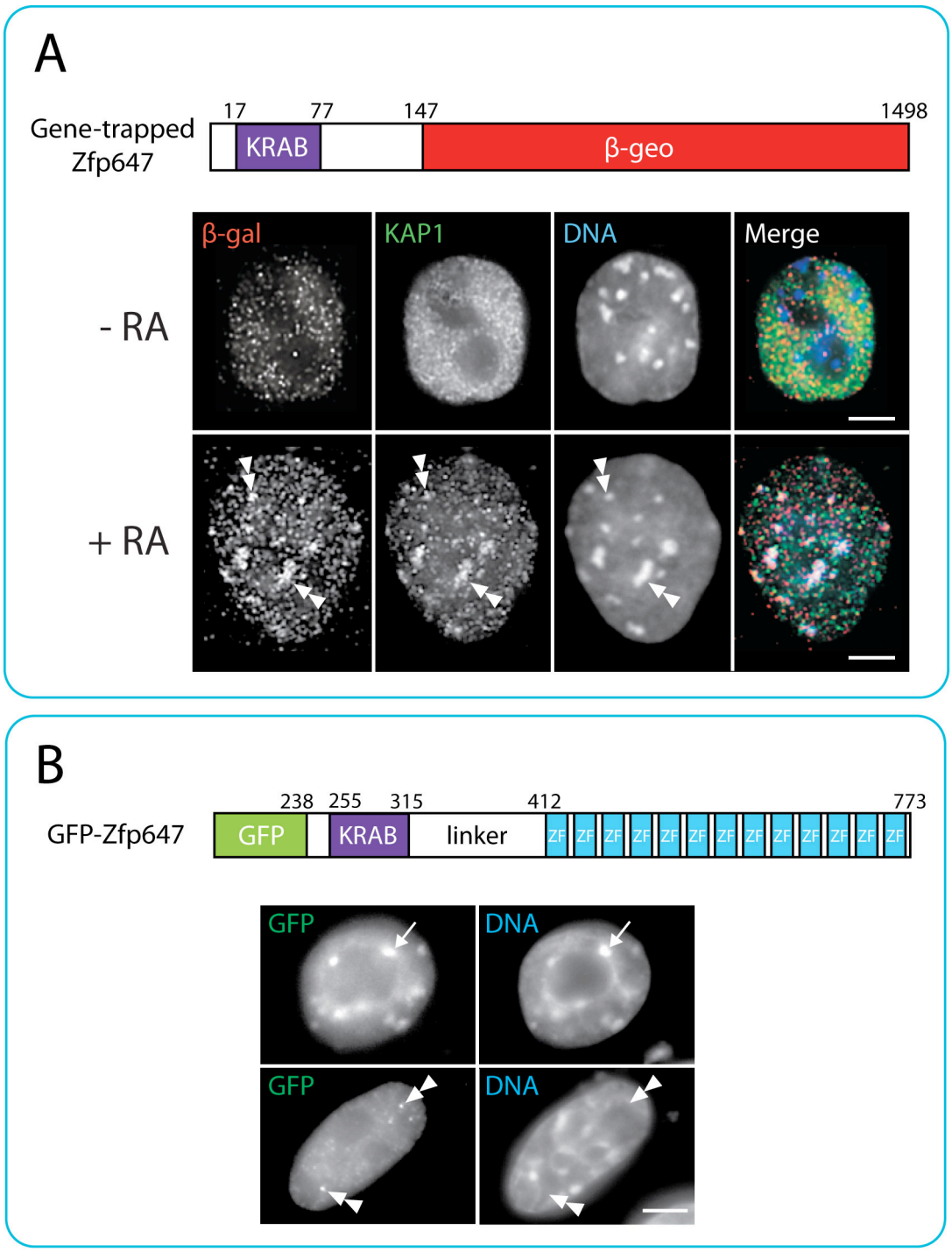


Figure 1.10. Localisation of gene-trapped and GFP-tagged Zfp647. A) Structure and localisation of gene-trapped Zfp647. Upper figure shows the gene-trap insertion site into Zfp647. Gene-trapped Zfp647 was detected by immunofluorescence with α-β-gal antibody in undifferentiated OS25 cells (-RA) and those differentiated with RA for ten days (+RA). KAP-1 was detected with mouse α-KAP-1 antibody. Double arrowheads indicate co-localisation of gene-trapped Zfp647 and KAP-1 at pericentric heterochromatin. B) Structure and localisation of GFP-Zfp647. Upper figure shows the N-terminally-tagged Zfp647 construct. Lower figure shows the localisation of GFP-Zfp647 and DNA in NIH/3T3 cells. Arrows indicate localisation of GFP-Zfp647 to the pericentric heterochromatin; double arrowheads indicate localisation of Zfp647 to non-heterochromatic foci. All immunofluorescence was carried out by Drs Heidi Sutherland and Stephanie Briers.

Briers against endogenous Zfp647 (one in sheep, one in rabbit) to verify the results obtained with the Zfp647 fusion proteins. The linker region (amino acids 90-174) was used as the antigen to immunise the animals, as this domain of the protein is the most unique (Figure 1.11A). The rabbit _-Zfp647 antibody stained within the nucleus of mouse cells, whereas the sheep _-Zfp647 antibody gave no signal by indirect immunofluorescence. Therefore all cell imaging experiments were carried out with rabbit _-Zfp647.

Co-staining of undifferentiated OS25 ES cells with the rabbit _-Zfp647 and mouse _-KAP-1 antibodies gave a similar fine speckled pattern of localisation for both proteins to that observed in the undifferentiated gene-trapped cells (Figure 1.11A, -RA). Differentiation of ES cells with RA dramatically changed the localisation of Zfp647 and KAP-1 in the nucleus. Zfp647 appeared less diffuse in the +RA cells and, instead of associating with the pericentric heterochromatin as found with the gene-trapped protein, the endogenous protein localised to relatively large, distinct non-heterochromatic foci (arrowed in Figure 1.11A, +RA). The number of cells displaying these foci increased throughout differentiation, reaching a peak of 44% of cells after 6 days of RA treatment, with between 5 and 40 foci present per nucleus (Briers *et al*, 2009). In differentiated cells, KAP-1 was strongly associated with the heterochromatic bright spots (double arrowheads in Figure 1.11), but also co-localised with Zfp647 at the non-heterochromatic foci, with at least two-thirds of the cells containing Zfp647 foci showing co-staining of KAP-1 (Figure 1.11A, +RA).

At this stage, it was unknown whether the relocalisation of Zfp647 to KAP-1 associated foci upon ES cell differentiation was a trait possessed by just this protein, or whether other KRAB zinc finger proteins could be found at these non-heterochromatic foci as well. Therefore an antibody was obtained from Yoshihiko Yamada which had been raised against another KRAB zinc finger protein, NT2 (Tanaka *et al*, 2002), and indirect immunofluorescence was carried out on undifferentiated and differentiated ES cells. This antibody was raised in rabbit, therefore co-staining could not be done with Zfp647. However, NT2 showed a similar pattern of localisation to Zfp647, with a fine speckled pattern of staining

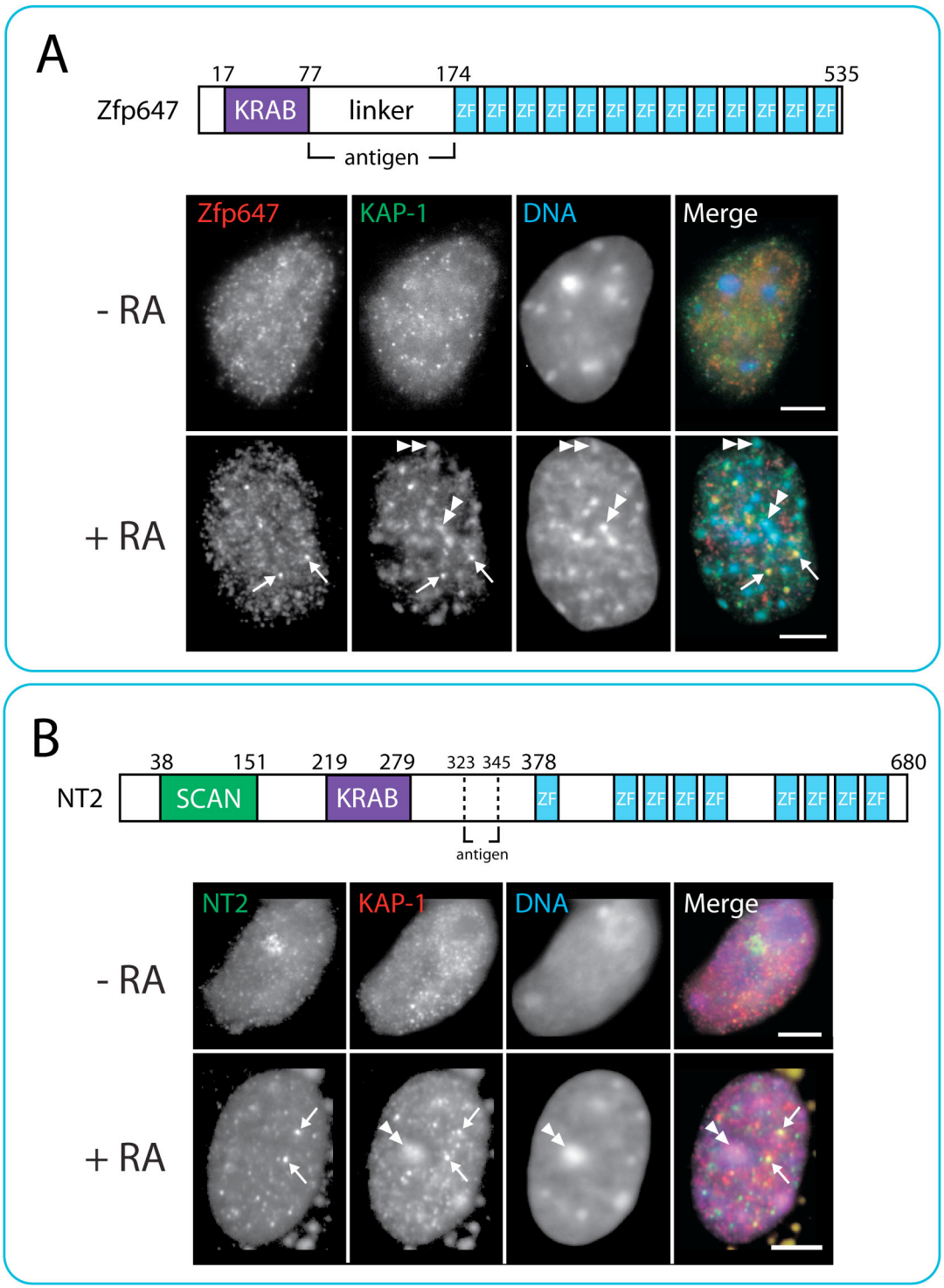


Figure 1.11. Localisation of KRAB zinc finger proteins in ES cells. A) Upper figure indicates the region of Zfp647 used as an antigen to generate α-Zfp647 antibodies. Lower figure shows localisation of Zfp647 (red) and KAP-1 (green) in undifferentiated (-RA) OS25 ES cells, and OS25 cells differentiated for 10 days with RA (+RA). Immunofluorescence was carried out with rabbit α-Zfp647 and mouse α-KAP-1 antibodies. B) Upper figure shows structure of NT2. Lower figure shows localisation of NT2 (green) and KAP-1 (red) in undifferentiated and differentiated OS25 cells. Immunofluorescence was carried out with rabbit α-NT2 (Tanaka *et al*, 2004) and mouse α-KAP-1 antibodies. Arrows indicate KRAB zinc finger protein co-localisation with KAP-1; double arrowheads indicate KAP-1 localisation to heterochromatin. All immunofluorescence was carried out by Drs Heidi Sutherland and Stephanie Briers.

reverting to punctate localisation at non-heterochromatic foci upon RA-mediated differentiation, co-incident with KAP-1 (Figure 1.11B) (Briers *et al*, 2009). We therefore believe that a number of KRAB zinc finger proteins may localize to these foci, and have thus named them KAKA foci, for KRAB and KAP-1 associated foci.

1.4.3. Co-localisation of Zfp647 with HP1 and PML

As discussed previously, the HP1 proteins play a critical role in the KRAB zinc finger protein transcriptional silencing mechanism, as loss of the interaction between KAP-1 and HP1 increases the expression of a KAP-1 target gene. If the KAKA foci formed a repressive environment within the nucleus where KRAB-mediated repression occurred, it was therefore possible that HP1 also localized to these foci. Incubation of differentiated OS25 cells with antibodies against HP1_ and HP1_ revealed that these proteins were often juxtaposed to the Zfp647 foci in the nucleoplasm, as well as localising to the pericentric heterochromatin (Figure 1.12A). The localization of another member of the KRAB repression machinery, SETDB1, was also analysed. This protein displayed a similar pattern of staining, often lying next to, or slightly overlapping, the KAP-1 foci (data not shown).

The KAKA foci appeared to be relatively large, protein-rich bodies. To determine whether there was any overlap with previously characterized nuclear bodies, co-staining with CREST antisera, which recognises centromeric components, was carried out with the α-Zfp647 antibody. Zfp647 did not localize to these structures in differentiated ES cells (data not shown). The average number of KAKA foci was similar to the number of PML bodies typically present in mammalian cells, and a number of transcription factors are known to be recruited to these bodies. Therefore co-staining studies were also carried out with Zfp647 and PML in RA-treated OS25 cells. These experiments showed that the Zfp647 foci often lay side-by-side with PML-NBs (Figure 1.12B) (Briers *et al*, 2009). This pattern of staining mirrors that of a human KRAB zinc finger protein, PAROT (PML-associated repressor of transcription), which localizes to foci within the nucleus alongside KAP-

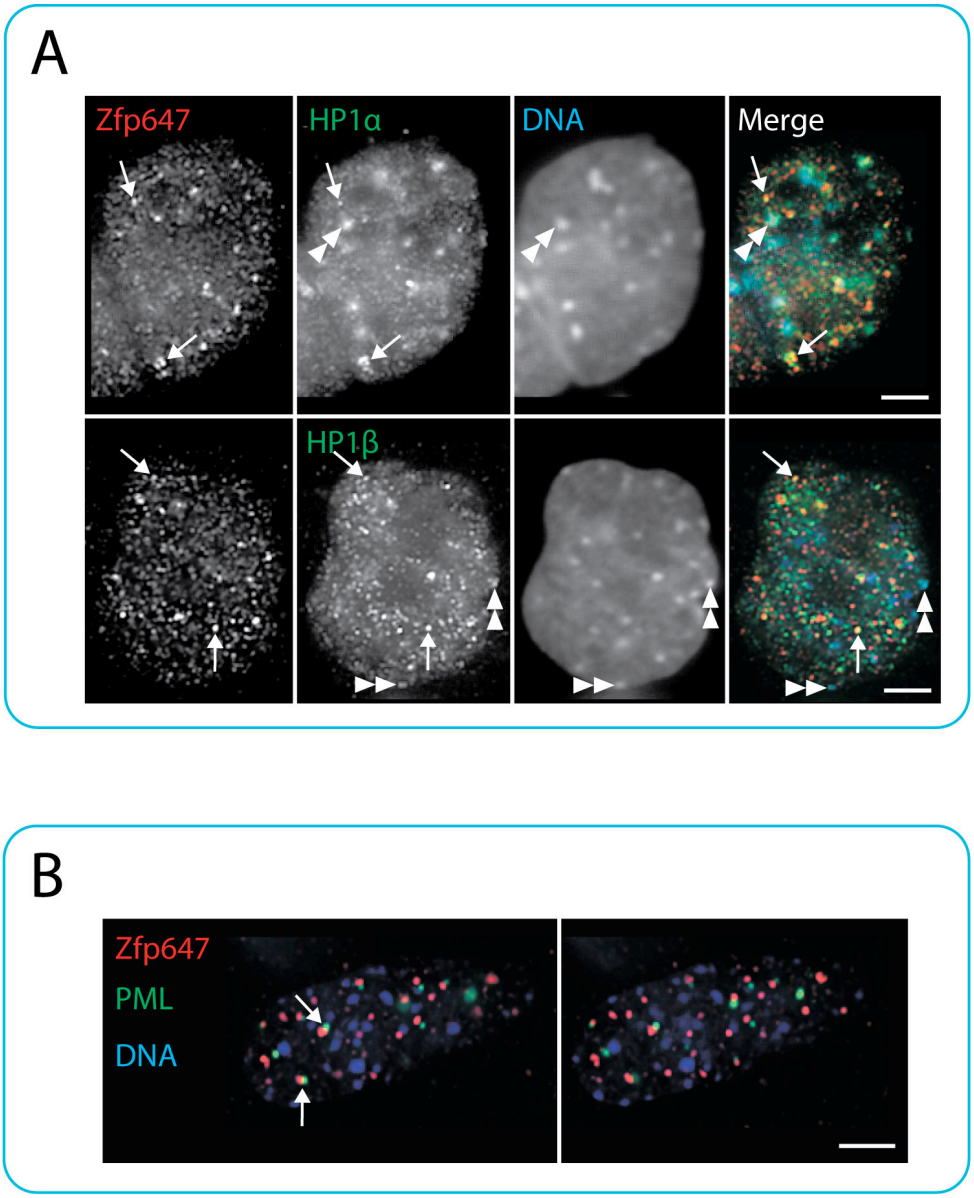


Figure 1.12. Localisation of Zfp647, HP1 isoforms and PML in differentiated ES cells. A) Localisation of Zfp647 (red) together with HP1α (upper panel) and HP1β (lower panel) (both green) in OS25 cells differentiated with RA for 10 days. Single arrowheads indicate co-localisation of Zfp647 and HP1; double arrowheads indicate co-localisation of HP1 and pericentric heterochromatin. B) Localisation of Zfp647 (red) and PML (green) in OS25 cells differentiated with RA for 10 days. Arrowheads indicate adjacent staining of the two proteins. All immunofluorescence was carried out by Dr Heidi Sutherland.

1, HP1_, HP1_, HP1_ and PML (Fleischer *et al*, 2006). However, these experiments were undertaken using a GFP-PAROT fusion protein, rather than immunofluorescence against the endogenous protein, thus it is possible that this protein is mislocalised, as found when the localization of endogenous and GFP-tagged Zfp647 were compared.

Why the KRAB zinc finger proteins relocalise to KAKA foci is currently unknown. We believe there are three main possibilities, which could overlap to some degree. First, the foci could be repressive environments, to which the KRAB zinc finger proteins and their gene targets are relocated during differentiation to cause gene silencing. The accumulation of many KRAB zinc finger proteins and a high concentration of KAP-1 may be beneficial to the repression process. Alternatively, the foci may be storage sites for KRAB zinc finger proteins and other KRAB-associated proteins, a function that has previously been suggested for the PML bodies, near to which the KAKA foci are found. Finally, given the proximity to the PML bodies, the KRAB zinc finger proteins may be recruited to the KAKA foci in order to undergo some form of post-translational modification, for example sumoylation (Vertegaal *et al*, 2006; Lee *et al*, 2007b; Mascle *et al*, 2007) or phosphorylation (Chang *et al*, 2007; Li *et al*, 2007).

1.5. Proposed Work

To discover more about Zfp647, as part of my Masters project, I undertook a yeast two-hybrid screen with the full-length Zfp647 protein and an embryonic mouse cDNA library. Protein-protein interactions between Zfp647 and transcriptional regulators of specific genes may suggest genes that Zfp647 helps to control, as has been shown with BRCA1 and ZBRK1 (Zheng *et al*, 2000), or potentially proteins that are present at the KAKA foci. During my Masters project, I managed to sequence each of the clones that I picked, however I did not verify any of these

interactions. Therefore my first aim for my PhD project was to test these proteins for their ability to auto-activate the nutritional markers used in the screen. It is also possible that some of these proteins interact with Zfp647 in yeast cells, but do not bind to Zfp647 *in vivo*. Thus I also wished to verify some of the interactions from the yeast two-hybrid screen *in vitro*.

One of the most interesting findings from the yeast two-hybrid screen was that Zfp647 is capable of interacting with a number of proteins involved in the ubiquitylation process. Given that KAP-1 has been shown to be sumoylated (Lee *et al*, 2007b; Mascle *et al*, 2007), and two other KRAB zinc finger proteins are also believed to be modified in this way (Vertegaal *et al*, 2006), I also set out to test whether Zfp647 is post-translationally modified, possibly by ubiquitin or another of the ubiquitin-like proteins. If this was the case, it may suggest that the KRAB zinc finger proteins, like KAP-1, are post-translationally modified to aid the transcriptional repression of their targets. Alternatively, modification may be important for the localisation pattern of Zfp647 seen in ES cells.

Chapter 2: Materials and Methods

2.1 Microbiological Methods

2.1.1. Bacterial Culture Conditions

All bacteria were cultured in Luria Bertani broth (LB), containing the antibiotic selective for the desired plasmid where appropriate (Table 2.1). BL21 (DE3) pLysS cultures were grown in the presence of chloramphenicol. Cultures were grown at 37°C unless otherwise stated.

Antibiotic	Stock Concentration	Working Concentration
Ampicillin	50mg/ml	50_g/ml
Chloramphenicol	34mg/ml	34_g/ml
Kanamycin	10mg/ml	30_g/ml

Table 2.1. Antibiotics used in bacterial culture.

2.1.2. Generating Competent Bacteria

1ml of a 3ml overnight BL21 (DE3) pLysS culture was used to inoculate a 100ml culture. Cells were left to grow at 30°C until an optical density at 600nm (OD_{600}) of 0.5-0.6 was reached. Cells were then pelleted at 6000g, resuspended in 25ml of 0.1M $CaCl_2$, and incubated on ice for 25 minutes. Cells were pelleted, resuspended in 10ml freezing mix (15% glycerol, 50mM $CaCl_2$), and divided into 200_l aliquots. Cells not to be used immediately were snap-frozen in dry ice and stored at -80°C until required.

2.1.3. Bacterial Transformation

Competent DH5_ (Invitrogen) or BL21 (DE3) pLysS cells were incubated with plasmid DNA on ice for 10 minutes. Cells were then heat-shocked at 42°C for 45 seconds, and returned to ice for a further 2 minutes. 300_l of LB was added, and the

cells were shaken at 37°C for 45 minutes. Transformed bacteria were plated on LB-agar plates containing the appropriate antibiotic and grown at 37°C overnight.

2.1.4. Making Extracts of Bacterially-Expressed Proteins

BL21 (DE3) pLysS cells transformed with pGEX-4T-1 carrying the protein of interest were grown in a 3ml overnight culture. The next morning, the bacteria were diluted into a 10ml culture to give an OD_{600} reading of 0.1-0.2, and the bacteria were grown at the appropriate temperature until the culture reached an OD_{600} value of 0.4. Two aliquots of 1ml were removed from the culture, and the remaining culture was induced to express the protein encoded on the transfected vector by the addition of isopropyl _-D-1-thiogalactopyranoside (IPTG) to 0.4mM. After 4 hours, two 1ml aliquots were removed from the culture. Total, soluble and insoluble cell extracts were made from the samples removed from the culture at both timepoints. To make a total cell extract, one of the 1ml samples was pelleted, and the pellet was weighed and resuspended according to its weight. The bacteria were resuspended in 3ml phosphate buffered saline (PBS) per gram of pellet, and an equal volume of 2x sodium dodecyl sulphate (SDS) loading buffer (100mM Tris-HCl (pH6.8), 4% (w/v) SDS (electrophoresis grade), 0.2% (w/v) bromophenol blue, 20% glycerol, 200mM _-mercaptoethanol) was added. To make soluble and insoluble extracts, the other sample was pelleted, and resuspended according to weight in lysis buffer (50mM Tris (pH8), 2mM ethylenediaminetetraacetic acid (EDTA) (pH8)): again, 3ml buffer was used per gram of pellet. 2_l 1% Nonidet P-40 (NP-40) and 2_l lysonase (Novagen) were added per 200_l of lysate, and the lysate was incubated at 37°C for 30 minutes. The insoluble fraction was pelleted by centrifugation in a Sorvall Benchtop microcentrifuge at 13000 rpm for 15 minutes at 4°C, and the supernatant was removed to a new tube. The insoluble matter was resuspended in the appropriate volume of lysis buffer or water, and an equal volume of 2x SDS loading buffer was added to both lysates. All samples were boiled at 95°C for 5 minutes. The solubility of the expressed protein was assessed by polyacrylamide gel electrophoresis (PAGE).

2.1.5. Refolding Insoluble Bacterial Proteins

The insoluble matter from bacteria expressing tagged proteins was pelleted as described above, weighed, and resuspended in dH_2O accordingly: for every 1g of pellet, 1ml dH_2O was used. All the following steps were carried out at 4°C. The insoluble lysate was divided into 40_l aliquots, re-pelleted for 15 minutes, and resuspended in 40_l 0.1M Tris (pH8.5) containing 0.5M, 1M, 2M or 5M urea. The insoluble proteins were spun down for 15 minutes, the supernatant was removed to a new tube, and the pellet was resuspended in 40_l dH_2O. An equal volume of 2x SDS protein loading buffer was added to the soluble and insoluble fractions, the samples were boiled at 95°C for 5 minutes, and analysed by PAGE.

2.2. Mammalian Cell Culture Methods

2.2.1. Mammalian Tissue Culture Conditions

NIH/3T3 and PCC4 cells were cultured in Dulbecco's-minimal Eagle's medium (D-MEM) (Cat #41965-039 Invitrogen) supplemented with 10% foetal calf serum (FCS), 100 units/ml penicillin and 100_g/ml streptomycin. Primary mouse embryonic fibroblasts (pMEF) cells were cultured in D-MEM supplemented with 15% FCS, 100 units/ml penicillin, 100_g/ml streptomycin, 0.3mg/ml l-glutamine, 1% MEM non-essential amino acids (Invitrogen) and 10_M _-mercaptoethanol. OS25 cells were cultured in Glasgow's minimal Eagle's medium (G-MEM) (Cat #21710-025) supplemented with 10% FCS, 100 units/ml penicillin, 100_g/ml streptomycin, 0.3mg/ml L-glutamine, 1% MEM non-essential amino acids, 1mM sodium pyruvate (Invitrogen) and 10_M _-mercaptoethanol on 0.1% gelatin-coated tissue culture flasks. To maintain OS25 cells in an undifferentiated state, they were additionally supplemented with 1000 units/ml human recombinant leukaemia inhibitory factor (LIF). All cells were grown at 37°C in 5% CO_2.

2.2.2. Differentiation of Tissue Culture Cells

To induce differentiation of OS25 cultures, cells were plated in media without LIF for one day. After this time, 5_M RA was added to the tissue culture media, and was included in the media for the remaining period of differentiation. Cells were passaged after two and four days of RA treatment. After four days of treatment with RA, 2.5_M gancyclovir was added to the culture media to select against any undifferentiated cells that were expressing Oct4.

PCC4 cells were induced to differentiate by the addition of 5_M RA into the culture medium. Differentiating cells were passaged as normal for six days.

2.2.3. Culturing pMEFs from Embryos

11.5dpc and 12.5dpc embryos were removed from their mother's uterus and placed in a 9cm tissue culture plate containing PBS supplemented with 100 units/ml penicillin and 100_g/ml streptomycin. The embryonic heads and internal organs were removed, and the remaining tissue of each embryo was transferred to one well of a 6-well plate containing 2ml of pMEF media. The following steps were carried out in a tissue culture hood. To separate the fibroblast cells, each embryo was placed into the barrel of a 5ml syringe and passed through a 21G needle 3 times. Cells were left to culture for 2-3 days, with the media being changed each day. When confluent, cells were trypsinised (50% trypsin in versene), resuspended in 5ml media and transferred to a 20ml universal tube. The largest particles of tissue were allowed to fall to the bottom, and the remaining cells in suspension were removed and cultured further.

2.2.4. Transfection of Tissue Culture Cells

Cells were seeded one day prior to transfection. Cells at 60-70% confluence were transfected using plasmid DNA and Lipofectamine 2000 (Invitrogen) according to the manufacturer's instructions. Cells were used 24 hours after transfection.

2.2.5. Freezing Mammalian Tissue Culture Cells

A confluent T75 or T175 tissue culture flask was trypsinised and resuspended in media. The cells were pelleted by centrifugation for 4 minutes at 1200rpm in a benchtop centrifuge, and washed once in PBS. The cells were pelleted again, and resuspended in freezing mix (10% dimethyl sulfoxide (DMSO) (Sigma) in FCS) according to volume of cells (1ml per $25cm^2$ of confluent cells). 1ml aliquots were transferred to Nunc Cryotube vials (Sigma), frozen down at -80°C, and transferred to liquid nitrogen for long-term storage.

2.3. Yeast Methods

2.3.1. Culture of Yeast Cells

Yeast cells were cultured in yeast peptone dextrose (YPD) or Sabouraud dextrose (SD) Broth lacking the appropriate nutrients, and shaken at 30°C. Yeast colonies were grown on YPD- or SD-agar plates lacking the appropriate nutrients at 30°C.

2.3.2. Rescue of Yeast Plasmids

To rescue yeast two-hybrid prey plasmids, the appropriate colony was picked and used to inoculate a 10ml culture of SD Broth supplemented with all nutrients except

leucine. The culture was grown for 2 days at 30°C, and the cells were pelleted in a bench top centrifuge at 3000rpm for 15 minutes. Cells were resuspended in 250_l P1 from the Qiagen Spin Miniprep kit, and transferred to a 1.5ml microcentrifuge tube containing 100_l of 400-600_m glass beads (Sigma). 4_l lyticase (Sigma) was added to the cells, which were incubated at 37°C for 1 hour. The cells were then smashed open by vortexing for 2 minutes, and the DNA was extracted by completing the miniprep kit protocol, including the optional washing step with buffer PB. DNA was eluted in 20_l dH_2O. The same day, 10_l of the extracted DNA was used to transfect 100_l of competent DH5_ cells (Invitrogen), which were grown overnight at 37°C. The DNA was then extracted from these cells by carrying out a miniprep on 3ml cultures of these cells.

2.3.3. Transformation of Yeast Cells

5ml of YPD (plus adenine) or the appropriate selective media was inoculated with a colony of the required yeast strain, and left to grow at 30°C for two days. 1ml of this culture was spun at 13000 rpm in a microcentrifuge tube for 30 seconds and the media was removed by inversion and shaking, leaving the cells in 50_l media. 2_l of salmon sperm carrier DNA (10mg/ml) was added, and the cells were resuspended. 1_g of the desired plasmid was added, and the cells were vortexed. 0.5ml of PLATE mixture (40.5% PEG, 0.1M lithium acetate, 10mM Tris-HCl (pH 7.5), 1mM EDTA) was added, and the cells were vortexed again. 20_l of 1M DTT was added, the cells were vortexed, and incubated on the bench overnight. The following day, the cells were heat-shocked for 10 minutes at 42°C, and spun down at 13 000 rpm for 30 seconds. The liquid was removed, the cells were washed in 0.5ml H_2O, re-pelleted, and resuspended in 100_l H_2O. 10_l of the cells, and the remaining cells, were plated onto the appropriate selective plates, and left to grow for 3-4 days at 30°C.

2.4. Preparation and Analysis of Proteins from Mammalian Cells

2.4.1. Whole Cell Protein Extracts

Cells in a confluent 9cm tissue culture plate were washed with PBS twice, and lysed in 500_l 2x SDS loading buffer. Lysates were sonicated with two 10-second pulses at 10_m on a Soniprep 150 sonicator (MSE), boiled at 95°C for 5 minutes and stored at -20°C. To preserve post-translational modifications of proteins, 20mM N-ethylmaleimide (NEM) was included in both the PBS washes and the 2x SDS loading buffer.

2.4.2. Nuclear and Cytoplasmic Extracts

Cells in a confluent 9cm tissue culture plate were washed twice with PBS, and scraped off in 500_l cell lysis buffer (10mM Tris (pH7.5), 100mM NaCl, 2.5mM $MgCl_2$, 35_g/ml digitonin, Complete protease inhibitors (Roche)). Cell lysates were incubated on ice for 5 minutes, and then centrifuged in a benchtop microcentrifuge at 4500rpm for 8 minutes at 4°C to separate the nuclear and cytoplamic fractions. The cytoplasmic supernatant was removed to a new tube, and the nuclear pellet was resuspended in 400_l nuclear lysis buffer (10mM Tris (pH7.5), 100mM NaCl, 2.5mM $MgCl_2$, 35_g/ml digitonin, 0.5% Triton X-100, Complete protease inhibitors (Roche)). The nuclear extract was incubated on ice for 5 minutes, sonicated twice for 10 seconds at 10_m, and cleared of insoluble material by centrifugation at 13000rpm for 5 minutes at 4°C in a benchtop microcentrifuge. If required for immediate analysis, an equal volume of 2x SDS loading buffer was added to the nuclear and cytoplasmic extracts, and the samples were boiled at 95°C for 5 minutes. To preserve post-translational modifications of proteins, 20mM NEM was included in the PBS washes, the lysis buffers and the 2x SDS loading buffer.

2.4.3. Polyacrylamide Gel Electrophoresis (PAGE)

6-12% polyacrylamide gels were cast according to Sambrook and Russell in Molecular Cloning (CSHL Press) using a Hoefer Inc. gel system. Protein samples and Benchmark pre-stained protein marker (Invitrogen) were loaded, and gels were run in Tris-glycine buffer (25mM Tris, 250mM glycine (pH8.3), 0.1% SDS). Alternatively, samples were run out on NuPAGE pre-cast gels (Invitrogen) according to the manufacturer's instructions. Gels were stained with Gel-Code Blue stain (Pierce) subjected to the western blotting procedure, or silver-stained.

2.4.4. Western Blotting

Proteins run out on polyacrylamide gels were transferred using a Genie Blotter (Idea Scientific) in buffer containing 25mM Tris, 190mM glycine and 20% methanol onto Hybond-P membrane (Amersham Biosciences) for one hour. Membranes were washed briefly in PBS containing 0.1% Tween-20 (PBST) and blocked in 10% western blocking reagent (WBR) (Roche) in PBST for one hour at room temperature or overnight at 4°C. Membranes were incubated in PBST containing 5% WBR and the appropriate dilution of primary antibody (see Table 2.2) for one hour at room temperature or overnight at 4°C. Blots were washed three times for 5 minutes in PBST, and incubated with a dilution of 1:10000-20000 secondary antibody-horseradish peroxidase (HRP) conjugate in 5% WBR in PBST for one hour at room temperature. Blots were washed five times in PBST, incubated in SuperSignal West Pico Chemiluminescent Substrate (Pierce) for 1 minute and exposed to autoradiography film (Amersham).

Antibody	Species	Dilution	Source (Product No.)
_-Zfp647	Sheep	1:1000 – 2000	Own
_-Zfp647	Rabbit	1:1000	Own
_-Rbak serum	Guinea pig	1:2500	Own
Affinity purified _-Rbak	Guinea pig	1:200	Own
_-KAP-1	Rabbit	1:2000	Bethyl Labs (A300-274A)
-Tif1	Mouse	1:2000	Gift from Prof Pierre Chambon (Weber *et al*, 2002)
_-GST	Rabbit	1:2000	Sigma (G7781)
_-His_6	Mouse	1:1000	Qiagen (34670)
_-c-myc	Mouse	1:250	Santa Cruz (sc-40)
_-HA	Rabbit	1:1000	Sigma (H6908)
_-GMP-1	Mouse	1:1000	Invitrogen (33-2400)
_-Sentrin-2	Rabbit	1:1000	Invitrogen (51-9100)
_-Ubiquitin	Rabbit	1:2000 – 5000	DakoCytomation (Z0458)
_-Nedd8	Rabbit	1:2000 – 5000	Abcam (ab4751)
_-tubulin	Mouse	1:5000	Sigma (T6793)
_-RanGAP1	Goat	1:2000	Abcam (ab2081)
_-Poly(ADP-ribose)	Mouse	1:500 – 1000	Abcam (ab14459)
_-rabbit secondary	Goat	1:10000	Sigma (A0545)
_-sheep secondary	Donkey	1:10000	Sigma (A3415)
_-mouse secondary	Goat	1:10000 - 20000	Sigma (A2304)
_-Guinea pig secondary	Rabbit	1:10000	Sigma (A5545)

Table 2.2. Antibodies used for western blotting and immunoprecipitations.

2.4.5. Immunoprecipitation of Mammalian Proteins

2.4.5.1. Conjugation of Antibody to Sepharose Beads

Protein G Sepharose 4 Fast Flow beads (GE Healthcare) were washed three times with 0.1M sodium phosphate buffer (pH 8.1). 20_l of sepharose was used per immunoprecipitation, and for each control. Antibody (see Table 2.2) was added at 0.2mg per ml of sepharose in a final volume of 0.7ml 0.1M sodium phosphate buffer, and the beads were left on a rotating wheel at 4°C for an hour. Beads were then washed in the buffer used to prepare the extracts to be immunoprecipitated, so as to remove any unbound antibody.

2.4.5.2. Immunoprecipitation from Cell Extracts

For small-scale immunoprecipitations, nuclear extracts were prepared as described above from 1x 9cm tissue culture plate. For large-scale immunoprecipitations for mass spectrometry experiments, nuclear extracts were prepared from 8x 14cm tissue culture plates. Nuclear extracts were pre-cleared using sepharose beads. 20-100_l protein G Sepharose 4 Fast Flow beads (GE Healthcare) were washed three times in nuclear lysis buffer (10mM Tris (pH7.5), 100mM NaCl, 2.5mM $MgCl_2$, 35_g/ml digitonin, 0.5% Triton X-100, Complete protease inhibitors (Roche)) and incubated in the extracts to be used for the immunoprecipitation for 1-4 hours at 4°C. The beads were then removed from the extracts by centrifugation, and the extracts were incubated with 20-100_l sepharose beads bound by the antibody of choice on a rotating wheel for between 1 hour and overnight at 4°C. Small-scale immunoprecipitations were carried out in a final volume of 700_l; large-scale immunoprecipitations were carried out in a final volume of 5ml. The beads were pelleted and washed up to six times for 10 minutes with 800_l of nuclear lysis buffer, each wash being carried out at 4°C. All remaining buffer was removed from the beads using a syringe attached to a 25G needle. The beads were resuspended in 25-50_l 2x SDS loading buffer, boiled at 95°C for 5 minutes, cooled on ice, and analysed by PAGE and western blotting or silver staining. Immunoprecipitations were stored at -20°C.

2.4.5.3. Silver Staining of PAGE Gels

Gels were fixed in 30% ethanol and 10% acetic acid for 30 minutes, followed by another fixation for 30 minutes in 30% ethanol, 0.5M sodium acetate, 0.2% soldium thiosulphate and 0.125% glutaraldehyde (glutaraldehyde was omitted from silver stained gels being analysed by mass spectrometry). The gels were then washed three times in dH_2O for 10 minutes, and stained by incubation in 0.1% $AgNO_3$, 0.01% formaldehyde for 20 minutes. The stain was removed, and the gels were rinsed briefly in a small amount of developer (2.5% $Na_2CO_3.10H_2O$, 0.01% formaldehyde),

which was aspirated off immediately. This process was repeated for as long as a precipitate was visible in the developer. Once the precipitate was no longer visible, the gels were incubated in the remainder of the developer solution until the gel was adequately stained. The staining was stopped by removal of the developer and incubation in 1% glycine until the gel was needed. To generate samples for mass spectrometry, gels were transferred to a 14cm tissue culture plate in a tissue culture hood, and the bands were excised using a clean razor blade. Samples were transferred to eppendorffs previously washed with methanol and distilled water, and samples were frozen at -20°C. For mass spectrometry analysis, the samples were transferred on ice to the SIRCAMS department at the University of Edinburgh.

2.4.6. Purification of His_6-Tagged Proteins

A 9cm tissue culture plate with cells at 60-70% confluence was transfected with plasmid DNA encoding His_6-tagged SUMO-1 (a gift from Professor Anne Dejean (Institut Pasteur, Paris) Müller *et al*, 2000), SUMO-2 (a gift from Professor Ron Hay (University of Dundee)), Ubiquitin (a gift from Dr Lee Finlan), Nedd8 (a gift from Dr David Girdwood), Zfp647 or the pcDNA3.1(+) vector. 24 hours later, the cells were scraped from the plate in 5ml media, and pelleted by centrifugation at 1200rpm for 5 minutes in a benchtop centrifuge. The cells were resuspended in 10ml PBS; 1ml was removed, the cells pelleted, and resuspended in 100_l 2x SDS loading buffer to allow the analysis of the input. The cells in the remaining 9ml sample were re-pelleted, and lysed in 5ml Buffer 1 (6M guanidium-HCl, 0.1M Na_2HPO_4, 0.1M NaH_2PO_4, 10mM Tris (pH8), pH to 8). At this point, cell lysates could be frozen at -80°C. Imidazole was added to the samples to a final concentration of 5mM, _-mercaptoethanol was added to a final concentration of 10mM, and the lysates were sonicated briefly. Any insoluble matter was removed by centrifugation at 4000rpm for 5 minutes at room temperature in a benchtop centrifuge, and the supernatant was transferred into a 15ml tube containing 50_l Ni^{2+}-NTA-agarose beads (QIAGEN) previously prepared by washing twice with 4ml Buffer 1 and pelleting at 2000rpm for 2 minutes. The cell lysates and agarose beads were left to rotate for 2-3 hours at

room temperature and then washed as follows: twice in 4ml Buffer 1 containing 10mM _-mercaptoethanol; three times in 4ml Buffer 2 (8M Urea, 0.1M Na_2HPO_4, 0.1M NaH_2PO_4, 10mM Tris (pH8), pH to 8) containing 10mM _-mercaptoethanol; twice in 4ml Buffer 3 (8M Urea, 0.1M Na_2HPO_4, 0.1M NaH_2PO_4, 10mM Tris (pH6.3) pH to 6.3) containing 10mM _-mercaptoethanol. The beads were transferred to a 1.5ml siliconized microcentrifuge tube in a final volume of 1.5ml Buffer 3 and pelleted at 2000rpm for 1 minute. Beads were then washed as follows: 1ml Buffer 3 containing 0.2% Triton X-100; 1ml Buffer 3; 1ml Buffer 3 containing 0.1% Triton X-100; three times in 1ml Buffer 3. The proteins were eluted from the beads by addition of 40_l elution buffer (for 1ml, mix 250_l 1M imidazole, 330_l 2x SDS loading buffer and 420_l dH_2O), mixed by finger-flick, and incubated either at 37°C for 1-2 hours, or overnight at 4°C. Beads and eluates were stored at -20°C, and proteins were analysed by PAGE and western blotting.

2.5. Generation of Guinea Pig _-Rbak Antibody

2.5.1. Large-Scale Expression and Purification of GST-Fused Rbak Protein

DNA encoding the linker region of Rbak was ligated into pGEX-4T-1 (GE Healthcare). This vector was transformed into competent BL21 (DE3) pLysS *E. coli* cells. Successfully transformed cells were used to inoculate a 20ml overnight culture. These cells were then diluted into a 400ml culture, which was left to grow at 37°C until it reached an OD_{600} of 0.6. The culture was induced by the addition of IPTG to 0.4mM, and was grown for a further 4 hours. Bacterial cells were then pelleted by centrifugation at 6000g at 4°C for 20 minutes, and were frozen down at -20°C overnight. The following day, cell pellets were resuspended in 15ml lysis buffer (20mM Tris (pH7.6), 1mM EDTA, 1mM ethylene glycol tetraacetic acid (EGTA), 100mM NaCl, 10% glycerol, 1mM dithiothreitol (DTT) and Complete protease inhibitors (Roche)). Lysozyme (Sigma) was added to 1mg/ml, and the cells were

incubated on ice for 30 minutes. The lysate was sonicated on ice for 5 minutes at 10_m, followed by the addition of Triton X-100 to 0.1% to ensure release of the fusion protein from the cells. The lysate was centrifuged at 12000g for 20 minutes at 4°C, and 15ml supernatant was removed. 2ml of 50% (v/v) glutathione-agarose (Sigma) previously suspended in PBS was added, and this was incubated at 4°C for 30-60 minutes on a roller mixer. The resin was loaded into a column (Econo, Bio-Rad), and the supernatant was allowed to drip through. A 1ml sample was retained as binding buffer (BB). The column was washed with 10ml PBST (PBS containing 1% Triton X-100 and protease inhibitors (Roche)): a 1ml sample was retained as wash buffer (W). Glutathione-S-transferase (GST)-Rbak was eluted by the addition of elution buffer (10mM reduced glutathione (Sigma) in 50mM Tris-HCl (pH8) (prepared fresh) plus Complete protease inhibitors (Roche)): five 1ml aliquots were passed through the column and collected separately as eluates 1-5 (E1-5). 50_l of each sample was removed, mixed with 50_l 2x SDS loading buffer and boiled at 95°C for analysis by PAGE and western blotting with _-GST antibody. After analysis, fractions E1 and E2 were pooled and concentrated in a centricon-10 column (Amicon/Millipore) following the manufacturer's instructions. All samples were frozen down at -20°C.

2.5.2. Immunisation of Guinea Pigs

250_g of purified GST-Rbak linker antigen was sent to Eurogentec, and was used to immunise two Guinea pigs using Eurogentec's standard 87-day immunisation programme. Animals were injected with 30_g antigen on days 0, 14, 28 and 56 of the programme, and were bled on days 0 (pre-immune), 38, 66 and 87 (final bleed). The immune responses of the animals were checked by western blotting NIH/3T3 extracts and bacterially expressed GST-Rbak linker protein with diluted pre-immune and final bleed sera.

2.5.3. Generation of Rbak Affinity Column

1g of cyanogen bromide-activated Sepharose 4B (Sigma) was resuspended in 3.5ml wash buffer (1mM HCl) and left to swell. The swollen gel was washed in a total of 200ml wash buffer using a glass scintered funnel, collected in a 15ml falcon tube, and made into a 50% slurry with wash buffer. For a 1.5ml column, 7.5mg of purified GST-Rbak linker protein (see above) was passed over a PD-10 column (GE Healthcare) previously equilibrated with coupling buffer (0.1M $NaHCO_3$, 0.5M NaCl, pH to 8), and eluted in 3.5ml coupling buffer. 3ml of the sepharose gel slurry was transferred to a clean 15ml falcon tube, and washed with 12ml of coupling buffer. The sepharose was centrifuged at 500g for 1 minute, the buffer removed, and the sepharose was resuspended in the antigen solution. The sepharose was mixed by inversion for 90 minutes (1 hour per ml of sepharose) at room temperature. Following this, the matrix was washed in 13ml of coupling buffer briefly, and then washed for 2 hours with 13ml of 0.1M Tris-HCl (pH8) at room temperature. The sepharose was washed with 0.1M sodium acetate (pH4), 0.5M NaCl for ten minutes, followed by one wash with 0.1M Tris-HCl (pH8), 0.5M NaCl, also for ten minutes. These two wash steps were repeated a further two times, followed by three ten minute washes with 13ml PBS. The matrix was resuspended in PBS, and poured into a column (Econo, Bio-Rad). The sepharose was washed with 15ml PBS, and the column was sealed with parafilm, leaving 3ml buffer on top of the sepharose, and stored at 4°C.

2.5.4. Affinity Purification of Polyclonal Rbak Antibodies

To purify the Guinea pig Rbak antibodies, serum was initially passed over a GST-LEDGF (lens epithelium-derived growth factor) column to remove antibodies against GST protein, and then purified on a GST-Rbak column. 5ml of Guinea Pig final bleed serum was diluted to 50ml in PBS, and passed through a 0.45_m filter. 10 bed volumes of PBS (i.e. 10ml for a 1ml column) were passed over a GST-LEDGF (p52) column previously made by Dr Heidi Sutherland, and the diluted serum was

then passed over this column three times. The GST-LEDGF column was subsequently washed, eluted and re-washed in the same manner as the GST-Rbak column to remove any antibodies against GST remaining on the column. The Guinea pig serum was then passed over the GST-Rbak affinity column three times, and the column was washed with 10 bed volumes of PBS. The column was then washed with 20 bed volumes of high salt wash buffer (10mM Tris-HCl (pH7.5), 500mM NaCl) followed by 20-50 bed volumes of PBS. Whilst washing the column, 10-15 microcentrifuge tubes were set up containing 35_l 1M Tris-HCl (pH8.8). When finished washing, 0.5ml aliquots of elution buffer (100mM glycine, 150mM NaCl, 10% glycerol) were allowed to pass over the column, collected in the microcentrifuge tubes, mixed by inversion immediately and kept on ice. The antibody concentration in each fraction was evaluated by diluting 5_l in 100_l PBS and measuring its UV absorbance value at 280nm (A_{280}) on a spectrophotometer. An A_{280} value of 1 equates to 0.75mg/ml of pure immunoglobulin G (IgG). The peak fractions were pooled and concentrated using a centricon-10 column (Amicon, Millipore), and the antibody was stored at -20°C.

2.6. Nucleic Acid Manipulation Methods

2.6.1. Gel Electrophoresis

1-2% agarose (Hi-Pure Low EEO Agarose, Biogene UK) gels were prepared using 1x TAE buffer (40mM Tris-acetate, 1mM EDTA (pH8)) or TBE buffer (45mM Tris-borate, 1mM EDTA (pH8)). Gels were run in 1x TAE or TBE containing 0.5mg/ml ethidium bromide. DNA samples, digests and polymerase chain reaction (PCR) products were loaded with 6x loading dye (Promega) and the gels were run at 5V/cm.

2.6.2. Plasmid DNA

2.6.2.1. Extraction of Plasmid DNA from Bacterial Cultures

For small-scale plasmid preps, 3ml of LB was inoculated with a single bacterial colony and grown overnight at 37°C. The next day, plasmid DNA was extracted using a QIAGEN QIAprep spin miniprep kit according to the manufacturer's instructions.

For large-scale preps, a 3ml LB culture containing a single bacterial colony grown over the course of a day was used to inoculate a 400ml culture which was grown overnight at 37°C. The next day, plasmid DNA was extracted using a QIAGEN Maxi prep kit according to the manufacturer's instructions, until the DNA was precipitated from the isopropanol solution. Following this, the DNA was resuspended in 500_l dH_2O, and transferred to a microcentrifuge tube. 50_l of 3M sodium acetate (pH5.5) was added, followed by 1ml of 100% ethanol, and the solution was inverted to precipitate the DNA. The DNA was pelleted by centrifugation at 13000rpm in a benchtop microcentrifuge at 4°C for 15 minutes, washed with 1ml 70% ethanol and re-centrifuged for 10 minutes. The DNA pellet was left to air-dry and resuspended in dH_2O according to the pellet size (typically 200_l).

2.6.2.2. Digestion of Plasmid DNA

Digestion of plasmid DNA was carried out using SuRE/Cut restriction enzymes (Roche) and the supplied buffer according to the manufacturer's instructions. DNA was analysed by gel electrophoresis, and where required for ligation, digests were purified by gel extraction.

2.6.2.3. Gel Extraction of Digested Plasmid DNA

Following gel electrophoresis, DNA samples were excised from the gel using a clean razor blade on a UV-light box. DNA was then extracted from the gel slice using the QIAGEN Gel Extraction kit. The DNA was typically eluted in 50_l dH_2O.

2.6.2.4. Ligation of Digested Samples

200ng of DNA, typically consisting of 100ng of each of the digested vector and insert sequences, 2_l 10 x ligation buffer (NEB), 1_l 1% NP-40, 2_l T4 DNA ligase (NEB) and 7_l H_2O were left at 16°C overnight. The following day, 10_l of the ligation was transfected into competent DH5_ cells.

2.6.3. Polymerase Chain Reaction

2.6.3.1. Sequencing of Plasmid DNA

500ng-1_g of plasmid DNA was mixed with 1_l of T7 sequencing primer (5': TAATACGACTCACTATAGGG) at 3.2_M and 4_l BigDye (Applied Biosystems) in a final reaction of 20_l. PCR was carried out on a DNA Engine Tetrad (MJ Research), for 25 cycles of 96°C for 30 seconds, 50°C for 15 seconds, and the appropriate annealing temperature for 4 minutes. To precipitate the DNA, 5_l of 125mM EDTA and 60_l 100% ethanol were added, and samples were incubated on ice for at least 20 minutes. The DNA was pelleted at 13000rpm for 30 minutes at 4°C, the pellet was washed in 70% ethanol and centrifuged at 13 000rpm for 15 minutes at 4°C, and was air-dried and handed in to sequencing.

2.6.3.2. Cloning of Plasmid DNA by PCR

20ng of plasmid DNA encoding the desired sequence was used in a 50_l PCR reaction, containing 1_l of forward and reverse primers at a concentration of 10_M, 1_l of dNTPs (Promega) at 10mM, 1_l of Turbo Pfu Polymerase (Stratagene), 5_l of 10x Pfu polymerase buffer and dH_2O up to 50_l. Table 2.3. shows the primer sequences used for generating each of the vectors described in the thesis. 5_l of the PCR products were analysed by gel electrophoresis. The remaining 45_l of PCR products were purified using the QIAquick PCR Purification Kit (QIAGEN) and eluted in 50_l of elution buffer. The PCR products and vectors to be cloned into were digested and ligated as described above, and transformed into competent DH5_ *E. coli* cells. Correct cloning was checked by sequencing of the newly created vectors.

Construct	Vector	Tag and Insert	Restriction Sites	5' and 3' Primers
pGEX-Zfp647	pGEX-4T-1 (GE Healthcare)	GST-Zfp647	5' – EcoRI 3' – XhoI	5' – ACTGAATTCGCAGCAGCCGGGCTTCTACC 3' – ACTCTCGAGGGGTCACAGCTTCTTGTGTAC
pGEX-Zfp647-ZF	pGEX-4T-1 (GE Healthcare)	GST-Zfp647	5' – EcoRI 3' – XhoI	5' – ACTGAATTCAGACCCTATATCTGCATTGAG 3' – ACTCTCGAGGGGTCACAGCTTCTTGTGTAC
pCGT7-Zfp647	pCGT7 (Cáceres *et al*, 1997)	T7-Zfp647	5' – XbaI 3' – BamHI	5' – ACTTCTAGAGCAGCAGCCGGGCTTCTACC 3' - ACTAGATCTGGGTCACAGCTTCTTGTGTAC
pCGT7-Zfp647-ZF	pCGT7 (Cáceres *et al*, 1997)	T7-Zfp647 zinc fingers	5' – XbaI 3' – BamHI	5' – ACTTCTAGAAGACCCTATATCTGCATTGAG 3' - ACTAGATCTGGGTCACAGCTTCTTGTGTAC
pcDNA3.1-Zfp647	pcDNA3.1(+) (Invitrogen)	His_6-Zfp647	5' – EcoRI 3' – XhoI	5' – ACTGAATTCATGCATCATCACCATCACCA TGCAGCAGCCGGGCTTCTACCG 3' – ACTCTCGAGCATGGGTCACAGCTTCTTGT GTACTTTCTG
pGADT7-Zfp647	pGADT7 (Clontech)	HA-Zfp647	5' – EcoRI 3' – XhoI	5' – ACTGAATTCGCAGCAGCCGGGCTTCTACCG 3' – ACTCTCGAGGGGTCACAGCTTCTTGTGTAC
pGBKT7-Zfp647	pGBKT7 (Clontech)	c-myc-Zfp647	5' – EcoRI 3' – SalI	5' – ACTGAATTCGCAGCAGCCGGGCTTCTACCG 3' - ACTGTCGACGGGTCACAGCTTCTTGTGTAC
pGEX-4T-1-Rbak linker	pGEX-4T-1 (GE Healthcare)	GST-Rbak linker	5' – EcoRI 3' – XhoI	5' – ACTGAATTCGTGGAAGGTGACCGCCATGCT 3' – ACTCTCGAGAGGGCTTCAGTTCCATTTGTG
pET-Rbakl linker	pET32a (Novagen)	His_6-Rbak linker	5' – EcoRI 3' – NotI	5' – ACTGAATTCGTGGAAGGTGACCGCCATGCT 3' – TGAGCGGCCGCAGGGCTTCAGTTCCAT TTGTG

Table 2.3. Plasmid constructs. Vectors and primers used to create the plasmids constructed as part of this thesis. NB: pCGT7 inserts were digested with BamHI; pCGT7 was digested with BglII.

2.7. *In Vitro* Expression of Proteins

2.7.1. *In Vitro* Transcription and Translation of Tagged Proteins

Target proteins were cloned into either of the two yeast two-hybrid vectors, pGADT7 or pGBKT7. Polymerase chain reaction (PCR) was carried out using primers designed to anneal from 6 nucleotides upstream of the T7 promoter, and at the stop codon of the inserted protein. The primer sequences used are shown in Table 2.4. PCR products were checked by agarose gel electrophoresis. 5_l of the 50_l reaction was used as the template in an *in vitro* transcription and translation (TNT) reaction using the TNT Quick for PCR DNA kit (Promega) according to the manufacturer's instructions. To radioactively label the protein, the reaction was supplemented with 4_l of [^{35}S]-methionine (1,000Ci/mmol at 10mCi/ml). The reaction was incubated at 30°C for 60-90 minutes, and stored at -20°C.

Primer	Sequence
pGADT7-Zfp647 5'	GAGATCTTTAATACGACTCACTATAGGGC
pGADT7-Zfp647 3'	ACTCTCGAGGGGTCACAGCTTCTTGTGTAC
pGBKT7-Zfp647 5'	GAATTTGTAATACGACTCACTATAGGGCG
pGBKT7-Zfp647 3'	ACTGTCGACGGGTCACAGCTTCTTGTGTAC

Table 2.4. Primers used to PCR-amplify tagged Zfp647 for *in vitro* expression.

2.7.2. Analysis of [^{35}S]-Labelled Proteins

To analyse [^{35}S]-labelled proteins, 2_l of a radioactive TNT reaction was added to 18_l of 2x SDS loading buffer and incubated at 95°C for 2 minutes. This was then loaded onto a pre-cast 10% Novex Bis-Tris gel (Invitrogen), and run according to the manufacturer's instructions. The gel was fixed in a solution containing 10% methanol, 10% acetic acid and 10% glycerol for 15 minutes. The gel was washed in dH_2O for 5 minutes, and then incubated with KODAK™ ENLIGHTNING™ Rapid Autoradiography Enhancer (Perkin Elmer) containing 10% glycerol for 30 minutes.

The gel was then dried under a vacuum at 80°C for 90 minutes, and exposed to autoradiography film (GE Healthcare) for 2-5 days.

2.7.3. Immunoprecipitation of [^{35}S]-Labelled Proteins

To analyse the efficiency of the immunoprecipitation of one protein, 45_l of an *in vitro* TNT reaction was incubated with sepharose beads which had been previously bound by the appropriate antibody, and 450_l immunoprecipitation (IP) buffer (20mM Tris (pH7.5), 100mM NaCl, 1% Triton X-100, 20% glycerol, Complete protease inhibitors (Roche)) for 90 minutes at 4°C on a rotating wheel. The beads were then washed three times with IP buffer, and resuspended in 20_l 2x SDS loading buffer. 10-20_l of the final volume was analysed by PAGE.

2.7.4. Co-Immunoprecipitation of *In Vitro*-Expressed Proteins

Two differently tagged proteins were *in vitro* transcribed and translated, one using unlabelled methionine contained within the kit, and the other with [^{35}S]-labelled methionine. After removing a 5_l sample for analysis, the remaining 45_l of each of the reactions were mixed together and incubated at 37°C for 1 hour. The co-immunoprecipitations were then carried out using the same method as an immunoprecipitation detailed above.

2.8. Fluorescent Methods

2.8.1. Indirect Immunofluorescence

Slides were seeded with cells one day prior to use. The following day, slides were washed twice for five minutes in PBS containing 1.5mM $MgCl_2$ and 1mM

$CaCl_2$ (all PBS used after this point also contained $MgCl_2$ and $CaCl_2$). Cells were fixed in 4% paraformaldehyde (pFA) in PBS for 20 minutes at room temperature, and then washed three times in PBS. Autofluorescence caused by pFA was quenched by treating the cells with PBS containing 50mM NH_4Cl for 10 minutes at room temperature. Cells were permeabilised with PBS containing 0.2% Triton X-100 for 12 minutes at room temperature, and then washed three times in PBS. Cells were blocked by incubation in 50_l PBS containing 5% donkey serum (Vector Labs) under a parafilm coverslip in a humidified chamber for 20 minutes at room temperature. Slides were then incubated under a fresh coverslip with 50_l blocking solution containing primary antibody (Table 2.5) for 1 hour at 37°C, or overnight at 4°C. Cells were washed three times in PBS, re-blocked for 10 minutes at room temperature, and then incubated in 50_l block containing the appropriate secondary antibody for 1 hour at 37°C. Slides were finally washed three times in PBS, mounted in Vectashield (Vector Labs) containing 0.4mg/ml DAPI under a glass coverslip and sealed with rubber solution (Truflex/Pang).

Antibody	Species	Concentration	Source (Product no.)
Zfp647	Rabbit	1:75	Own
KAP-1	Rabbit	1:200	Bethyl Labs (A300-274A)
Tif1_	Mouse	1:200	Gift from Prof Pierre Chambon (Weber *et al*, 2002)
PML	Mouse	1:200	Millipore (MAB3738)
Zfp37	Rabbit	1:100	Gift from Dr Niels Galjart (Payen *et al*, 1998)
Zfp90	Rabbit	1:100	Gift from Dr Niels Galjart (Payen *et al*, 1998)
Uhrf1	Rabbit	1:200	Gift from Dr Pier Paolo Di Fiore (Citterio *et al*, 2004)
Affinity purified Rbak	Guinea pig	1:20-100	Own
FITC-_-rabbit secondary	Goat	1:100	Vector Labs (FI-1000)
Texas red-_-rabbit secondary	Goat	1:100	Vector Labs (TI-1000)
FITC-_-mouse secondary	Horse	1:100	Vector Labs (FI-2000)
Texas red-_-mouse secondary	Horse	1:100	Vector Labs (TI-2000)
FITC-_-Guinea pig secondary	Goat	1:100	Vector Labs (FI-7000)

Table 2.5. Antibodies used for indirect immunofluorescence experiments.

Chapter 3: Protein-Protein Interactions of Zfp647

3.1. Introduction

During my Masters project, I carried out a yeast two-hybrid screen to identify potential Zfp647-interacting proteins. Zfp647 (as well as another KRAB zinc finger protein, NT2) has been shown to localise to foci containing KAP-1 and HP1 in differentiated ES cells, which are also situated adjacent to the PML-NBs. To identify other proteins that might be in KAKA foci, or other proteins involved in the function of Zfp647, I used the full-length sequence of Zfp647 as the bait. Dr Heidi Sutherland and Dr Stephanie Briers have studied the expression of Zfp647 in mouse embryos and adults, and found that the protein is expressed from at least 8.5 days post-coitum (dpc) until 15.5dpc, thus I used an 11.5dpc mouse embryonic cDNA library as the prey for my screen. The yeast two-hybrid screen, undertaken as part of my Masters project in the Bickmore lab, yielded 107 potential Zfp647-interacting proteins; however, none of these interactions were verified during the course of the project. I therefore sought to test the results by using the yeast two-hybrid method to identify false positive results. Following this, I intended to try and co-immunoprecipitate endogenous Zfp647-interacting proteins from tissue culture cells, and also prove certain interactions made by Zfp647 using tagged, exogenously expressed proteins.

3.2. Verification of Yeast Two-Hybrid Results

3.2.1. Yeast Two-Hybrid Screen Verification Method

The potential Zfp647 interactors identified in my yeast two-hybrid screen fell into six main categories: KRAB zinc finger and tripartite motif proteins (Table 3.1); non-KRAB domain-containing zinc finger proteins (Table 3.2); chromatin-associated proteins (Table 3.3); proteins involved in post-translational modification (Table 3.4); RNA-binding and translation proteins (Table 3.5); and cytoskeletal proteins

Protein Name (Accession No.)	No. of Hits in Screen	Minimal Amino Acid Startpoint	Strength of Interaction	Clone Used in Auto-Activation	Auto-Activation
Zfp647 (NM_172817)	30	231	++	10	No
Zfp1 (NM_001037665)	1	208	++	193	No
Zfp26 (NM_011753)	4	129	+/++	83	No
Zfp30 (NM_013705)	2	232	+/++	191	No
Zfp51 (NM_009558)	1	494	+++	90	No
Zfp60 (NM_009560)	1	402	++	123	No
Zfp68 (NM_013844)	2	245	++/+++	320	No (poor growing)
Zfp75 (NM_172918)	1	389	++	336	No
Zfp90 (NM_011764)	2	365	++/+++	464	No
Zfp101 (NM_009542)	1	206	+	46	No
Zfp169 (NM_026450)	2	119	++	39	No
Zfp189 (NM_145547)	1	313	++++	226	No
Zfp202 (NM_030713)	1	223	+	450	No
Zfp354c (NM_013922)	1	265	++	316	No
Zfp386 (NM_019565)	1	350	+++	199	No
Zfp563 (NM_001024950)	1	84	++	669	No
Zfp612 (NM_175480)	13	267	+++	145	No
Rslcan2 (NM_001001130)	1	73	++	452	No
Rslcan11 (NM_001012448)	2	218	++/+++	136	No
Rslcan15 (NM_177712)	1	94	+++	641	No
Rbak (NM_021326)	1	386	++	533	No
PREDICTED: RIKEN cDNA 4930422I07 gene (XM_150216)	1	244	++	261	No
RIKEN cDNA 2210010B09 gene (NM_172919)	1	341	+	365	No
Trim23 (NM_030731)	1	8	++++	387	No
Trim28 (KAP-1) (NM_011588)	31	119	+++	2	No

Table 3.1. Auto-activation results for the KRAB zinc finger and tripartite motif proteins rescued from Zfp647 yeast two-hybrid screen. Strength of interaction (+ to ++++) was estimated by X-_-gal staining.

Protein Name (Accession No.)	No. of Hits in Screen	Minimal Amino Acid Startpoint	Strength of Interaction	Clone Used in Auto-Activation	Auto-Activation
Zfp3 (NM_177565)	2	164	+++/++++	250	No
Zfp62 (NM_009562)	2	625	+++	164	No
Zfp280c (NM_153532)	2	574	++	289	No
Zfp319 (NM_024467)	3	303	++++	241	No
Zfp623 (NM_030199)	7	117	+++/++++	54	No
Zfp668 (NM_146259)	6	168	+++	71	No
Zfp768 (NM_146202)	1	334	+++	177	No
Zfp771 (NM_177362)	7	127	++	16	No
Gli3 (NM_008130)	1	1	++	352	Partial
Zscan21 (NM_001044703)	1	122	+	453	No
Zscan22 (NM_001001447)	1	326	+++	651	Yes
PREDICTED: RIKEN cDNA 4930461P20 gene (Dnajc21) (XM_001473643)	3	202	++++	95	Yes

Table 3.2. Auto-activation results for zinc finger proteins rescued from Zfp647 yeast two-hybrid screen. Strength of interaction (+ to ++++) was estimated by X-_-gal staining.

Protein Name (Accession No.)	No. of Hits in Screen	Minimal Amino Acid Startpoint	Strength of Interaction	Clone Used in Auto-Activation	Auto-Activation
Prdm5 (NM_027547)	1	335	+	76	No
Prdm15 (XM_907577)	3	402	++	141	No
Chromodomain (similar to HP1_) (NM_007624))	1	1	+	178	No
PREDICTED: Jmjd1c (Jumonji domain containing 1c) (XM_907708)	1	1494	+++	282	No
PREDICTED: Baz1a (XM_885873)	2	1296	++/+++	310	No
Uhrf1 (NM_010931)	3	503	++++	326	No
Morf4l1 (Mortality factor 4 like 1) (NM_024431)	1	6	+	746	No
Snw1 (NM_025507)	1	1	+	754	No

Table 3.3. Auto-activation results for chromatin binding proteins rescued from Zfp647 yeast two-hybrid screen. Strength of interaction (+ to ++++) was estimated by X-_-gal staining.

Protein Name (Accession No.)	No. of Hits in Screen	Minimal Amino Acid Startpoint	Strength of Interaction	Clone Used in Auto-Activation	Auto-Activation
Mnat1 (Ménage a trois 1) (NM_008612)	1	15	++++	198	No
Usp33 (Ubiquitin-specific peptidase 33) (NM_133247)	2	761	++++	284	No
Uba52 (Ubiquitin A-52 residue ribosomal protein fusion product 1) (NM_019883)	19	18	++	312	No
Ube3a (ubiquitin protein ligase E3A) (NM_011668)	2	667	++++	757	Yes
PREDICTED: Ubc (Ubiquitin C) (XM_001479832)	1	32	++++	682	Yes

Table 3.4. Auto-activation results for ubiquitin machinery proteins rescued from Zfp647 yeast two-hybrid screen. Strength of interaction (+ to ++++) was estimated by X-_-gal staining.

Protein Name (Accession No.)	No. of Hits in Screen	Minimal Amino Acid Startpoint	Strength of Interaction	Clone Used in Auto-Activation	Auto-Activation
Rps7 (Ribosomal protein S7) (NM_011300)	81	1	+	91	No
Rps5 (Ribosomal protein S5) (NM_009095)	1	166	+	53	No
Hnrpll (Heterogeneous nuclear ribonucleoprotein L-like) (NM_144802)	1	56	+++	441	No
Eif3g (Eukaryotic translation initiation factor 3, subunit G) (NM_016876)	1	197	+	503	No
F1l1 (FIP1 like 1) (NM_024183)	1	427	++++	548	Yes
Rpl5 (Ribosomal protein L5) (NM_016980)	3	88	++	608	No
Rbm17 (RNA binding motif protein 17) (NM_152824)	1	163	+	681	Yes on 3x

Table 3.5. Auto-activation results for translation-associated proteins rescued from Zfp647 yeast two-hybrid screen. Strength of interaction (+ to ++++) was estimated by X-_-gal staining.

(Table 3.6). Proteins that did not belong to any of these groups are shown In Table 3.7. Many of these proteins have known DNA-binding domains such as zinc fingers, thus it is possible that a number of these may be false positive results. Zfp647 was cloned into the Clontech vector pGBKT7, which fuses the DNA binding domain of the GAL4 transcription factor to the bait protein. The mouse cDNA library had been cloned into the pGADT7 vector, which creates a fusion of the prey protein and the GAL4 activation domain. The yeast strain AH109, into which the bait vector was transformed, carries the nutritional markers *ADE2* and *HIS3*, which are under the control of GAL4 promoters. Y187 also carries the *lacZ* and *MEL1* genes, which encode the _-galactosidase and _-galactosidase enzymes. After mating the AH109 bait strain and the A187 prey yeast strain, an interaction between a prey protein and the bait leads to the coming together of the two GAL4 domains. This allows the expression of the *ADE2* and *HIS3* genes, enabling the mated yeast strain to survive on media deficient in adenine and histidine. The initial positive results of the screen were selected for in this way. An interaction between the bait and prey also leads to the expression of the *lacZ* and *MEL1* genes, which can catalyse the breakdown of the chemicals X-gal or X-_-gal releasing a dark blue pigment. The strength of the interaction between the bait and the prey proteins can then be assessed by the blueness of the yeast colony. However, it is possible that as the prey proteins are already fused to the GAL4 activation domain, if they bound to the promoters of the selective markers without interacting with the bait protein, the markers would be transcribed, giving a false positive result.

To identify any false positive clones from the screen, one clone encoding each interactor was rescued from the yeast colonies picked from the screening plates. These clones are shown in Tables 3.1 – 3.7, and were sequenced to verify the prey protein present. They were then transformed into yeast strain AH109 carrying either the empty pGBKT7 vector or pGBKT7-Zfp647. Any clone that allowed growth on media deficient in adenine and histidine in the absence of Zfp647 demonstrated that it did not require another protein's presence to express the nutritional markers, and was therefore a false positive result (Figure 3.1). 15 clones did not grow when re-cultured, therefore these proteins could not be tested for auto-activation. Of the 92

Protein Name (Accession No.)	No. of Hits in Screen	Minimal Amino Acid Startpoint	Strength of Interaction	Clone Used in Auto-Activation	Auto-Activation
Kif5a (Kinesin family member 5A) (NM_008447)	1	223	+	6	No
Rdx (Radixin) (NM_009041)	6	1	++	18	No
Becn1 (Beclin1) (NM_019584)	1	110	++++	48	No
Tbcb (Tubulin folding cofactor B) (NM_025548)	1	32	+++	156	Yes
Myh10 (NM_175260)	5	1636	++	188	No
Kif5b (Kinesin family member 5B) (NM_008448)	1	364	+	254	No
Spag5 (Sperm associated antigen 5) (NM_017407)	9	678	++	270	No
PREDICTED: Myh3 (Myosin, heavy polypeptide 3) (XM_908146)	2	1450	+/++	342	No
Ndel1 (Nuclear distribution gene E-like homolog 1) (NM_023668)	1	1	++++	430	No
Lasp1 (LIM and SH3 protein 1) (NM_010688)	1	1	+	445	No
Myo6 (Myosin VI) (NM_001039546)	1	978	++	535	No
Vim (Vimentin) (NM_011701)	1	10	+++	660	No

Table 3.6. Auto-activation results for cytoskeletal proteins rescued from Zfp647 yeast two-hybrid screen. Strength of interaction (+ to ++++) was estimated by X-_-gal staining.

Protein Name (Accession No.)	No. of Hits in Screen	Minimal Amino Acid Startpoint	Strength of Interaction	Clone Used in Auto-Activation	Auto-Activation
Coiled coil domain containing 136 (Ccdc136) (XM_485735)	10	71	++	19	No
Ran binding protein 9 (Ranbp9) (NM_019930)	4	49	++	192	Yes
Wilms' tumour 1-associating protein (Wtap) (NM_175394)	11	1	+++	203	No
Phosphatase and tensin homolog (Pten) (NM_008960)	1	1	+	252	No
MTERF domain containing 1 (Mterfd1) (NM_025547)	10	110	++/+++	280	No
ATPase, H^+ transporting, V0 subunit C (Atp6v0c) (NM_009729)	1	32	++	335	Yes
RIKEN cDNA 4930432O21 gene (XM_001471514)	1	233	++	381	No
Thioredoxin 2 (NM_019913)	2	1	++++	496	No
Myelocytomatosis oncogene (Myc) (NM_010849)	1	151	+	739	No
Replication factor C 1 (Rfc1) (NM_011258)	1	821	++	758	No

Table 3.7. Auto-activation results for unclassified proteins rescued from Zfp647 yeast two-hybrid screen. Strength of interaction (+ to ++++) was estimated by X-_-gal staining.

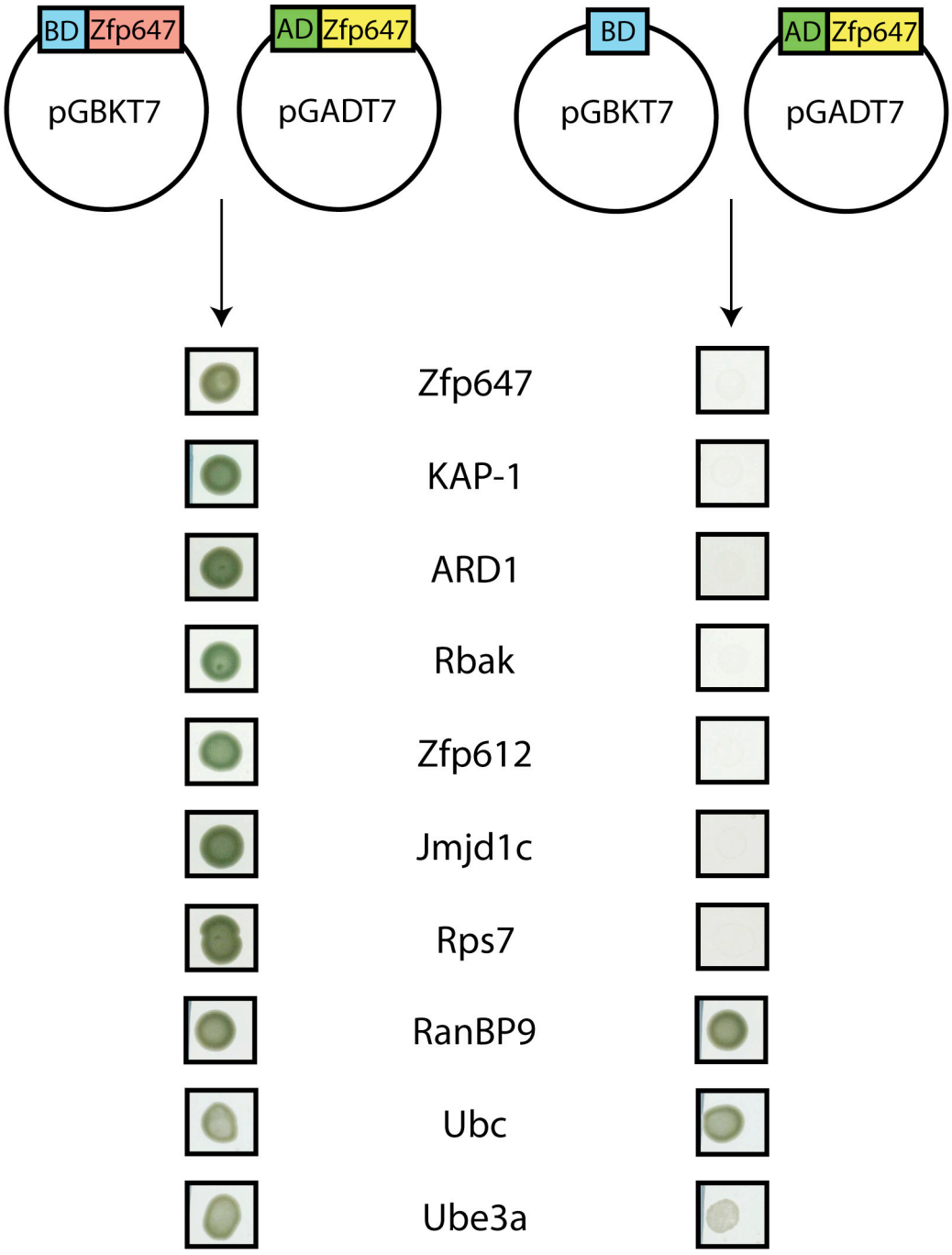


Figure 3.1. Yeast two-hybrid screen auto-activation test. Yeast cells were co-transfected with pGADT7 carrying Zfp647 and either the empty pGBKT7 vector (right side) or pGBKT7-Zfp647 (left side). Plasmids conferring growth on selective media in the absence of the bait Zfp647 protein were deemed false positives. Results are shown for 10 of the 92 proteins tested. BD: GAL4 DNA binding domain; AD: GAL4 activation domain

proteins that were tested, ten grew when co-transformed with the empty pGBKT7 vector, and were therefore false positives. These proteins were the GLI-Krüppel family member GLI3 (Gli3); zinc finger and SCAN domain containing 22 (Zscan22); DnaJ (Hsp40) homolog, subfamily C, member 21 (Dnajc21); ubiquitin protein ligase E3A (Ube3A); ubiquitin C (Ubc); RNA binding motif protein 17 (Rbm17); RAN binding protein 9 (Ranbp9); FIP 1 like 1 (F1l1); ATPase, H^+ transporting V0 subunit C (Atp6v0c); and tubulin folding cofactor B (Tbcb).

3.2.2. Interaction with KAP-1 and ARD1

Of the 800 clones recovered in the screen, 31 encoded the known KRAB zinc finger protein co-repressor protein KAP-1. This therefore acted as a positive control, demonstrating that Zfp647 was capable of interacting correctly in the yeast cells. Previous studies have shown that the interaction between KAP-1 and the KRAB domain is mediated by the amino terminus of KAP-1, and it is still unclear from the literature which domains of KAP-1 are required for this interaction. To identify KAP-1's interaction domain, I sequenced each of the KAP-1 yeast clones from the N-terminus. Many clones in a cDNA library are missing the 5' end of the mRNA sequence, due to 5'-3' degradation of the mRNA during the preparation of the library, thus the remainder of the sequence must contain the interacting domain. In the case of the KAP-1 clones, 31 positive colonies were picked and sequenced, and were found to encode 11 distinct clones. The shortest protein recovered started at amino acid 119, C-terminal of the RING domain (Figure 3.2). This indicates that the RING finger is not necessary for interacting with KRAB zinc finger proteins, and thus it is likely that only the B-boxes and coiled coil domains are required for KAP-1's interaction with Zfp647, as shown previously for KAP-1's interactions with the KRAB domains of ZNF10 and KRAZ1 (Moosmann *et al*, 1996; Agata *et al*, 1999).

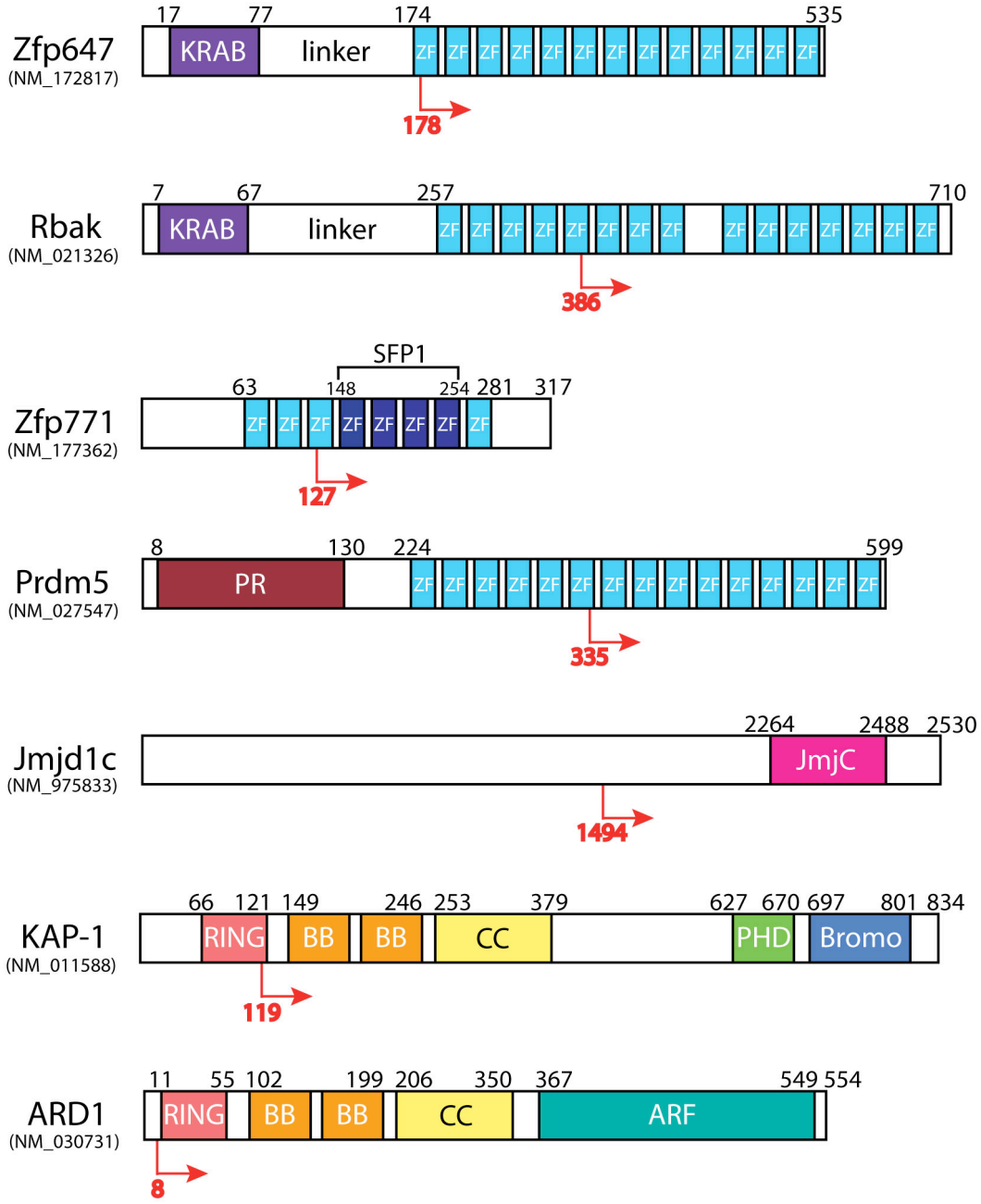


Figure 3.2. Zfp647-interacting proteins. Diagram showing domains of some of the zinc finger and tripartite motif proteins identified as interacting partners of Zfp647 in yeast two-hybrid screen. Numbers above the protein structure depict the amino acid at the start/end of each domain. The first amino acid of the protein coded in the recovered yeast two-hybrid prey vector is shown in red below the domain structure. KRAB: Krüppel-associated box; ZF: zinc finger; PR: positive regulatory domain; JmjC: jumonji; RING: really interesting new gene; BB: B-box; CC: coiled-coil; PHD: Plant homeodomain; Bromo: bromodomain; ARF: ADP-ribosylation factor.

Zfp647 also interacted with another tripartite motif protein known as Trim23 or ARD1 (ADP-ribosylation factor domain protein 1). It has been previously reported that the human homologue of Zfp647, ZNF250, interacts with the human ARD1 protein in a high throughput yeast two-hybrid screen (Rual *et al*, 2005). This screen utilized the human ORFeome v1.1 library, which encodes ~ 7200 distinct proteins. None of the other 13 ZNF250-interacting proteins found in the screen by Rual *et al* were identified in my Zfp647 yeast two-hybrid screen. Similarly to my results, ZNF250 was shown to interact with DNA-binding proteins (ZNF165, LDOC1), chromatin-binding proteins (ZBTB8) and cytoskeletal proteins (HOOK1, HOOK2 and KRTAP4-12) (Rual *et* al, 2005). ARD1 and KAP-1 both encode an RBCC domain at their amino terminus, and thus may also interact with Zfp647 through the same portion of this structure. Because the only ARD1 clone recovered from the screen initiated at the eighth amino acid of the protein and therefore also retained the RING motif (Figure 3.2), I could not determine whether this is in fact the case from this analysis.

Unlike the transcription intermediary factors Tif1_, Tif1_ (KAP-1), Tif1_ and Tif1_, ARD1 does not encode a PHD and bromodomain at its C-terminus; instead it has an ADP-ribosylation factor domain. Proteins with this domain are associated with vesicular transport, and can also stimulate ADP-ribosyltransferase activity mediated by the cholera toxin, although the only function that ARD1 has been shown to possess is *in vitro* ubiquitin E3 ligase activity (Vichi *et al*, 2005). No *in vivo* targets of this activity have been demonstrated to date (Professor Martha Vaughan (National Institutes of Health, Bethesda), personal communication). Antibodies have been raised against the protein, but the native protein has not been detected in any tissues or cells tested thus far (Professor Martha Vaughan (National Institutes of Health, Bethesda), personal communication).

3.2.3. KRAB Zinc-Finger Protein Homo- and Hetero-Oligomerisation

One of the most interesting findings from the yeast two-hybrid screen was the revelation that Zfp647 could interact with itself, and 23 other KRAB zinc-finger proteins. Zfp647 was encoded in 30 of the 800 clones sequenced, making it the third highest 'hit' in the yeast two-hybrid screen, after ribosomal protein S7 and KAP-1 (see Table 3.1). One other KRAB zinc-finger protein, ZBRK1 is known to homo-oligomerise through an extended C-terminal domain which is not present in Zfp647 (Tan *et al*, 2004b), thus another part of Zfp647 must be responsible for self-interaction. Sequencing of the 30 positive colonies from the screen revealed that 4 distinct clones were present. The cDNA sequences of these clones started at nucleotide positions 529, 530, 658 and 692 of the Zfp647 sequence; the shortest of these initiates at the second amino acid of Zfp647's third zinc finger (Figure 3.2). Therefore it is most probable that it is through the last 11 zinc fingers that Zfp647 binds to itself.

Each of the other KRAB zinc-finger proteins from the screen were analysed in the same manner, and it was found that the KRAB domain, and varying portions of the linker domain, was missing from each of these sequences. This suggests that the zinc fingers of Zfp647, as found with the KRAB zinc-finger proteins KZLP (Watanabe *et al*, 2007) and Nizp1 (Nielsen *et al*, 2004), have a protein-binding function in addition to their putative DNA-binding capabilities. Some of the proteins identified in the yeast two-hybrid screen (Zfp1, Zfp354c and Zfp612) show a relatively high level of homology to Zfp647 (between 66 and 70%) across their zinc finger domains. I therefore investigated whether the zinc finger domains of the KRAB zinc finger proteins identified in the screen were more similar to Zfp647 than a group of KRAB zinc finger proteins not identified in the screen. I chose to include some proteins related to those recovered from the screen (for example, Rslcan12 and Zfp354a) to see whether they were more dissimilar to Zfp647 than Rslcan11 and Kid-3. Philippe Gautier (MRC Human Genetics Unit) aligned the peptide sequences of the zinc finger domains using the MEGA program, an integrated software tool for Molecular Evolutionary Genetics Analysis (http://www.megasoftware.net). The zinc

finger domains had variable numbers of zinc finger motifs, and any gaps in the alignment were removed from the sequence in a pairwise fashion. The minimal bootstrap value for the phylogenetic tree produced (Figure 3.3) was 50%, meaning that any of the branches seen in the tree must be present in at least 50% of the trees generated by the program. No significant clustering of any of the KRAB zinc finger proteins was observed in the tree, and it appears that those proteins from the yeast two-hybrid screen (shown in red in Figure 3.3) were no more similar to Zfp647 than the other proteins (shown in green in Figure 3.3). These results suggest that the interactions between Zfp647 and KRAB zinc finger proteins identified by the yeast two-hybrid screen are not due to a high level of similarity between the proteins' zinc finger domains.

3.2.4. Zfp647 Interactions with Other Transcription Factors

Zfp647 interacted with 12 zinc finger proteins (Table 3.2). Three of these proteins (Gli3, Zscan22 and Dnajc21) auto-activated the expression of the yeast nutritional markers. Six of the proteins that did not autoactivate the markers (Zfp3, Zfp280c, Zfp319, Zfp623 and Zfp768) encoded only zinc finger motifs, making it extremely likely that Zfp647 interacted with these proteins through their zinc fingers. Disregarding the false positive results, the remaining proteins contained two different domains: Zfp62 and Zfp771 encode a SFP repressor domain, and Zscan21 contains a SCAN domain (Figure 3.2). Those proteins containing a SFP domain are believed to be transcriptionally repressive, whereas Zscan21 is thought to activate transcription. Thus Zfp647 may potentially interact with both transcriptional activators and repressors. Additionally, Zfp647 interacted with SNW domain containing 1 (Snw1), a protein that binds to Daxx, and that can function as a transcription co-factor (Tang *et al*, 2005). Snw1 has no recognisable domains when analysed by the SMART program (http://smart.embl-heidelberg.de/). When over-expressed in tissue culture cells, Snw1 co-localises with Daxx, often at the PML-NBs, to which Daxx is recruited by PML (Tang *et al*, 2005). This protein may therefore provide a link between the KAKA foci and the PML-NBs.

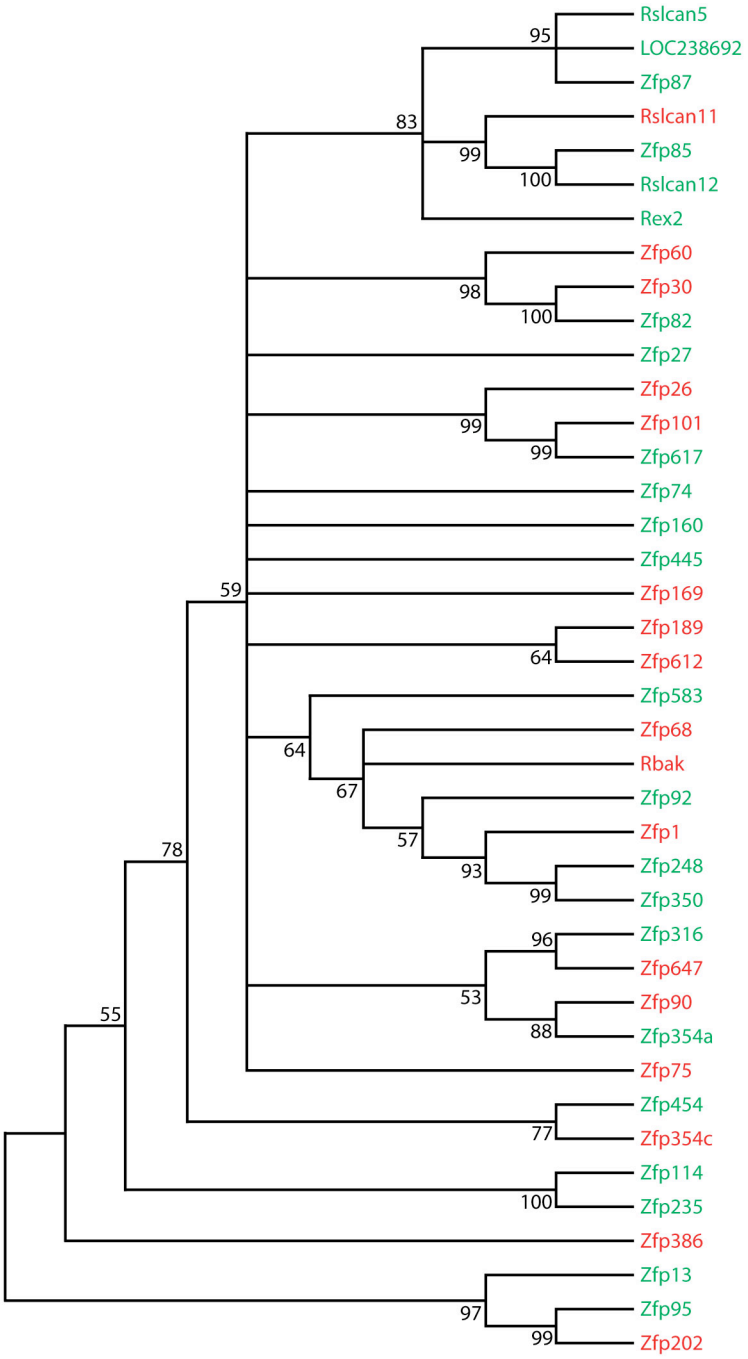


Figure 3.3. Alignment of KRAB zinc finger proteins. Phylogenetic tree representing the alignment of the amino acid sequences of the zinc finger domains of the 17 annotated KRAB zinc finger proteins identified in the Zfp647 yeast two-hybrid screen (shown in red), and 23 other KRAB zinc finger proteins (shown in green). Rslcan5, Rslcan12 and Zfp354a were chosen due to their relation to Rslcan11 and Zfp354c; other proteins were randomly selected. Alignment was carried out using the MEGA program: gaps in the sequences were removed in a pairwise fashion, and the bootstrap values (shown to the left of the branches) had a 50% cutoff. The alignment was carried out by Philippe Gautier.

3.2.5. Chromatin-Associated Proteins

None of the chromatin-associated proteins that were pulled out in the screen caused the autoactivation of the nutritional markers (Table 3.3). These proteins were PR domain containing 5 (Prdm5) and Prdm15, mortality factor 4 like 1 (Morf4l1), bromodomain adjacent to zinc finger domain 1A (Baz1a), jumonji domain containing 1c (Jmjd1c), ubiquitin-like, containing PHD and RING finger domains, 1 (Uhrf1) and an unknown chromodomain protein (Figures 3.2 and 3.4). Morf4l1 can associate with both HAT and HMTase complexes (Pardo *et al*, 2002), and loss of the function of this protein is lethal during embryogenesis (Tominaga *et al*, 2005). Uhrf1 is known to bind to and ubiquitylate histone proteins (Citterio *et al*, 2004). It is also responsible for loading Dnmt1 onto hemi-methylated DNA, and so is essential for the maintenance of DNA methylation (Sharif *et al*, 2007). By contrast, the other three proteins have not been studied to the same degree, although it is known that they encode domains capable of binding to, and possibly altering, histone modifications. Jmjd1c encodes a jumonji domain, which can catalyse the removal of methyl groups from histones, although Jmjd1c has to date not been found to possess demethylase activity. Similarly, the PR domain found in Prdm5 and Prdm15 is highly similar to the SET domain, which functions as a protein methyltransferase, but Prdm5 and Prdm15 have yet to be shown to be capable of methylating histones or other proteins. However, Prdm5 is known to interact with the repressive H3K9 HMTase G9a (Duan *et* al, 2007). Thus Zfp647 may interact with chromatin modifying complexes other than SETDB1 and Mi-2α.

3.2.6. Ubiquitin Machinery / RING Domain Proteins

Eight proteins that are involved in the ubiquitylation process, or that have the ubiquitin E3 ligase RING domain were pulled out of the yeast two-hybrid screen. Two of these proteins were false positives: Ube3a and Ubc. The five remaining proteins are ménage a trois 1 (Mnat1), ubiquitin specific peptidase 33 (Usp33), ubiquitin A-52 residue ribosomal protein fusion product 1 (Uba52), Uhrf1 and ARD1

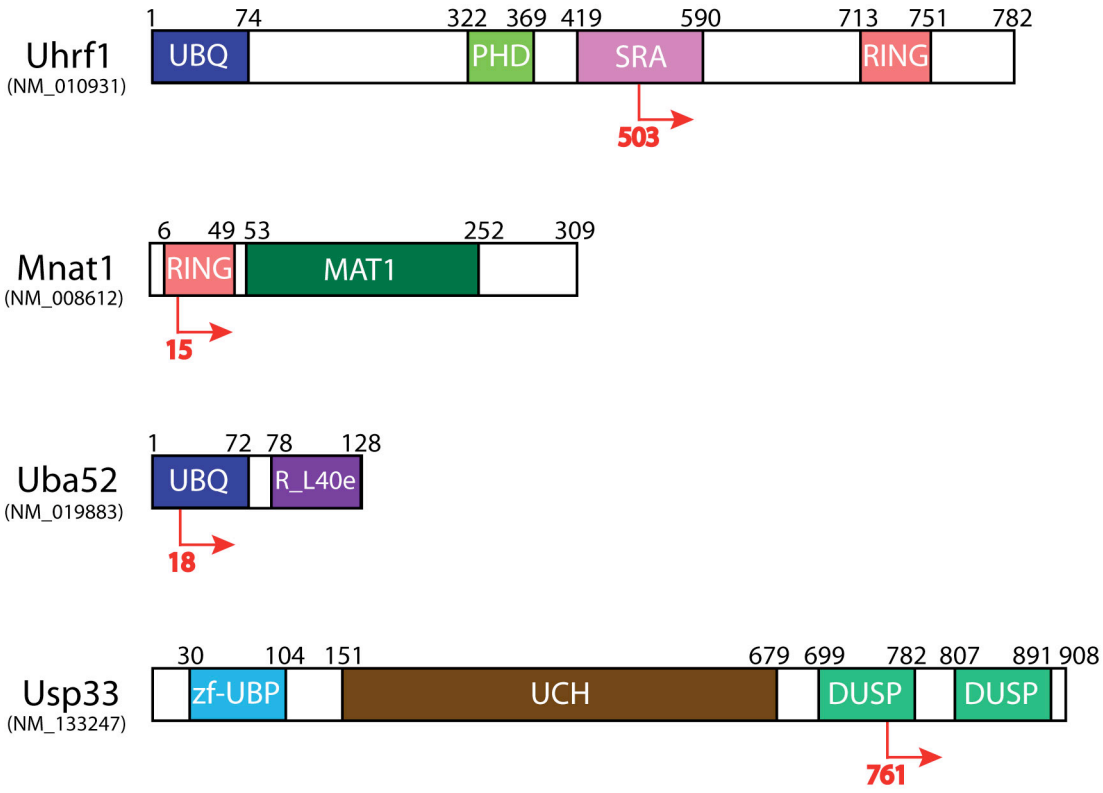


Figure 3.4. Ubiquitin machinery Zfp647-interacting proteins. Diagram showing domains of some of the RING domain-containing and ubiquitin machinery proteins identified as interacting partners of Zfp647 by yeast two-hybrid screen. Numbers above the protein structure depict the amino acid at the start/end of each domain. The first amino acid of the protein coded in the recovered yeast two-hybrid prey vector is shown in red below the domain structure. UBQ: ubiquitin homologue; SRA: SET and RING finger-associated domain; MAT1: menage a trois 1; R_L40e: ribosomal_L40e domain; zf-UBP: zinc finger-ubiquitin binding domain; UCH: ubiquitin C-terminal hydrolase; DUSP: domain in ubiquitin-specific peptidases.

(Figure 3.4 and Table 3.4). Mnat1, Uhrf1 and ARD1 have a RING domain; Uhrf1 and ARD1 possess ubiquitin E3 ligase activity. Uhrf1 can ubiquitinate histones in an *in vitro* context, in particular histone H3 (Citterio *et al*, 2004). ARD1 can ubiquitylate itself, GST protein and UbcH6 (an E2 component) when subjected to *in vitro* ubiquitylation assays (Vichi *et al*, 2005). Narrowing down the interactive domains of these proteins from the plasmid sequences, Zfp647 may bind to the RING domains of ARD1, Uhrf1 and Mnat1 (Figure 3.4). The ubiquitin machinery proteins that were not false positives encode a variety of domains: the MAT1 (Mnat1), DUSP (Usp33), UBQ and ribosomal_L40e domains (both Uba52) (Figure 3.4). These proteins have varying functions in ubiquitylation. Uba52 is a source of ubiquitin within the cell, whereas Usp33 removes ubiquitin molecules from poly-ubiquitylated substrates (Li *et al*, 2002). Mnat1 encodes a RING domain, making it a potential E3 ligase, but thus far is known to act as a stabilising factor to the Cdk7-cyclin H complex, which is responsible for phosphorylating serine 5 (Ser5) of the heptapeptide repeats found in the C-terminus of RNA polymerase II, allowing it to initiate transcription (Devault *et al*, 1995; Fisher *et al*, 1995; Tassan *et al*, 1995). It is unknown why Zfp647 might be interacting with these particular ubiquitin-related and RING domain proteins, but these results suggest that there is a link between Zfp647 and ubiquitylation.

3.2.7. RNA Binding and Translation Machinery Proteins

One RNA-binding protein (Rbm17) of the seven tested via the auto-activation method yielded a false positive result (Table 3.5). The protein recovered the most times from the yeast two-hybrid screen, ribosomal protein S7 (Rps7), did not autoactivate the marker genes. Although the ribosomes are mainly located in the cytoplasm, and Zfp647 is found in the nucleus, Rps7 could interact with Zfp647 in an endogenous setting, as it has been shown to be imported into the nucleus by four different import receptors (Jäkel and Görlich, 1998). Rps7 can interact with other non-ribosomal proteins, as it has been demonstrated to bind to MDM2, preventing the MDM2-mediated ubiquitylation and subsequent degradation of the tumour

suppressor protein p53 (Chen *et al*, 2007). Whether the interaction between Zfp647 and Rps7 does occur endogenously, however, requires further investigation.

3.2.8. Cytoskeletal Proteins

The twelve cytoskeletal proteins were also tested for their potential for auto-activation: one protein (Tbcb) auto-activated the nutritional markers (Table 3.6). Despite the number of different cytoskeletal proteins pulled out in the screen, I believe it is unlikely that these factors bind to Zfp647 endogenously, firstly because of their differing functions to that of Zfp647, and secondly because of the very poor growth of the colonies when they were re-tested. The majority of the proteins tested for auto-activation grew well on the selective media, whereas two-thirds of the cytoskeletal proteins did not grow on at least one of the two selective plates they were tested on. Therefore I believe these positive results were due to an interaction that only occurred in the yeast cells. Similarly, many of the proteins pulled out of the screen but not classified into a particular functional group, shown in Table 3.7, gave poor growth, and also localize to structures such as the mitochondria (e.g. Thioredoxin), thus it is unlikely that many of these proteins interact with Zfp647.

3.3. Cellular Localisation of KRAB Zinc-Finger Proteins and Potential Zfp647-Interacting Proteins

Studies on the sub-cellular localisation of various KRAB zinc-finger proteins have generally been carried out using the overexpression of a fluorescently tagged fusion protein, allowing visualisation of the protein without having to raise an antibody against it. Due to the high homology between KRAB zinc finger proteins, particularly those encoded in multi-gene clusters, raising specific antibodies against these proteins may be problematic. However, we have found that the localisation of a

GFP-tagged Zfp647 construct does not completely mirror that of the endogenous protein, revealed by immunofluorescence studies using a rabbit _-Zfp647 antibody (Briers *et al*, 2009), suggesting that an exogenous protein does not necessarily behave in the same manner as the endogenous protein. A number of the papers using this method of over-expression to study KRAB zinc finger proteins may have lost the detailed pattern of staining within the cell, and instead only achieve a general picture of localisation (Qi *et al*, 2005; Ou *et al*, 2005; Cao *et al*, 2005). A few studies have outlined more specific localisations of KRAB zinc finger proteins, in some cases co-staining with interacting proteins. For example, Lee *et al* have described an interaction between the zinc fingers of the human protein ZNF133 and the protein inhibitor of activated STAT1 (PIAS1), and demonstrated that the proteins co-localised at nuclear speckle structures by co-expression of the tagged proteins (Lee *et al*, 2007c). PAROT localises to punctate foci within the nucleus, and directly co-localises with KAP-1, HP1_, HP1_, HP1_ and PML (Fleischer *et al*, 2006).

Nevertheless, only a small number of KRAB zinc finger proteins have had antibodies raised against them. Tanaka *et al* raised a rabbit polyclonal antibody against the mouse KRAB zinc finger protein NT2, although they did not use this antibody to study the sub-cellular localisation of the protein (Tanaka *et al*, 2002). Huang *et al* used a rabbit polyclonal antibody against ZNF23 to detail its nuclear localization (Huang *et al*, 2007). Most pertinently for this study, Payen *et al*, as well as raising an antibody against Zfp37, which again shows nuclear localisation, raised an antibody against Zfp90, one of the KRAB zinc finger proteins found to interact with Zfp647 in my yeast two-hybrid screen (Payen *et al*, 1998). The localisation of Zfp90 was not outlined in this paper, however. Zfp90 is the only murine KRAB zinc-finger protein from the yeast two-hybrid screen to have an antibody raised against it to date. Another potential interactor that has had antibodies raised against it is Uhrf1, which binds to and ubiquitinates histones (Citterio *et al*, 2004).

3.3.1. Localisation of Zfp647 in Mouse Cells

Drs Heidi Sutherland and Stephanie Briers showed that Zfp647 can localize to non-heterochromatic foci in differentiated ES cells (Briers *et al*, 2009). However, in NIH/3T3 cells, the localization of Zfp647 was only analysed with the use of the GFP-tagged Zfp647 construct, which we believe mislocalises. I therefore decided to study the localisation of endogenous Zfp647 in two cell lines, NIH/3T3 cells and primary mouse embryonic fibroblasts pMEF cells. This would allow me to see whether Zfp647 exhibits the same behaviour in these cells as in mouse ES cells, and therefore whether KAKA foci are a common occurrence in different cell lines, or whether they are restricted to specific cell types.

NIH/3T3 cells were grown on glass slides, fixed in pFA and incubated with rabbit _-Zfp647, mouse _-KAP-1 and mouse _-PML antibodies (Section 2.8.1; Figure 3.5A-C). In NIH/3T3 cells, Zfp647 appears to be diffuse within the nucleus. The large punctate foci observed in ES cells did not appear to be present, although Zfp647 gave a slightly granular appearance, with some small foci present (arrowed in Figure 3.5B and C). However, Zfp647 was not seen to co-localise with KAP-1 in a significant number of nuclei, as seen in differentiated ES cells. Similarly, I did not see the twinning of these smaller Zfp647 foci with PML, which was seen in differentiated ES cells. KAP-1 is seen to localize to the pericentric heterochromatin in differentiated ES cells (Figure 1.11). I did not see this localisation in NIH/3T3 cells; instead, KAP-1 was more diffuse, with a granular staining pattern (Figure 3.5A and B).

I also studied the localisation of Zfp647 in pMEFs. I generated these cells from 12.5dpc mouse embryos (Section 2.2.3), and carried out immunofluorescence on early passage cells (either passage 5 or 6) with rabbit _-Zfp647, mouse _-KAP-1 and mouse _-PML antibodies (Figure 3.5D-E). In approximately 15% of the cells analysed, Zfp647 could be seen at distinct nuclear foci. These Zfp647 foci co-localised with KAP-1 (arrowed in Figure 3.5D), which could also be seen at the pericentric heterochromatin (double arrowheads in Figure 3.5D). Similar to the

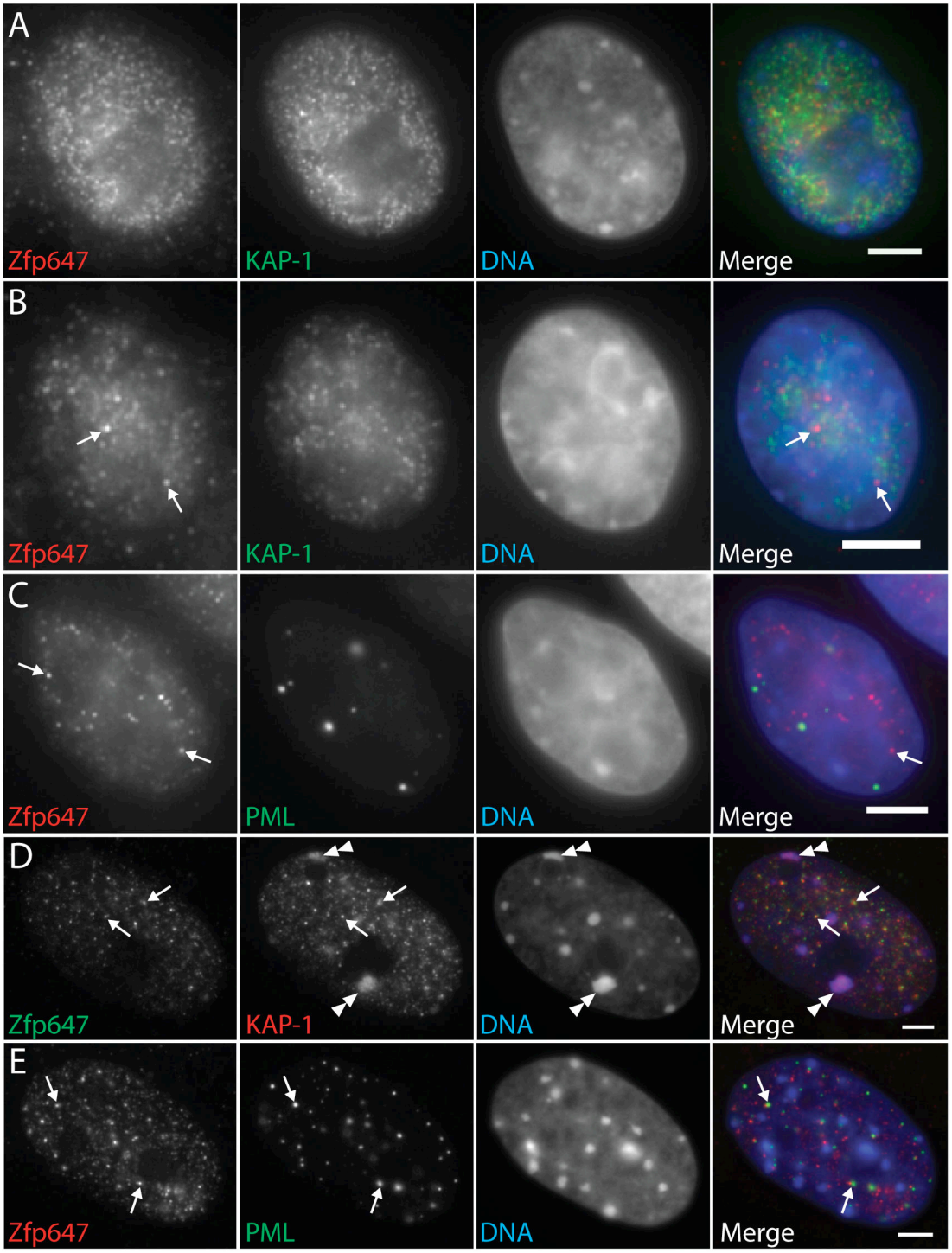


Figure 3.5. Localisation of Zfp647, KAP-1 and PML in NIH/3T3 and pMEF cells. Immunofluorescence was carried out using A) - C) NIH/3T3 cells or D) and E) pMEFs. A), B) and D) cells were incubated with rabbit α-Zfp647 and mouse α-KAP-1. C) and E) cells were incubated with rabbit α-Zfp647 and mouse α-PML. Signals were detected with Texas red-α-rabbit and FITC-α-mouse secondary antibodies. DNA was detected with DAPI. Scale bars represent 5μm. Arrows indicate putative Zfp647 foci; double arrowheads indicate KAP-1 localisation to the pericentric heterochromatin.

localisation results from differentiated ES cells, the Zfp647 foci were also adjacent to the PML nuclear bodies in approximately 10% of cells (double arrowheads in Figure 3.5E). Therefore while it appears that the KAKA foci are not obvious in NIH/3T3 tissue culture cells, they are present in a proportion of pMEFs.

3.3.2. Immunofluorescence of Potential Zfp647 Interacting Proteins

Two antibodies against other KRAB zinc finger proteins (Zfp37 and Zfp90) were received from Dr. Niels Galjart (Erasmus MC, Rotterdam). Both of these antibodies were raised against the linker regions of the proteins, and both were raised in rabbits, therefore it was not possible to co-stain nuclei for any of the KRAB zinc finger proteins together. I co-stained NIH/3T3 cells with the antibodies specific for either Zfp37 or Zfp90 with mouse _-KAP-1. Dr Heidi Sutherland performed immunofluorescence on differentiated ES cells using the same antibodies. The immunofluorescence results for the NIH/3T3 cells are shown in Figure 3.6A-D. Zfp37 appeared to be relatively diffuse throughout the nucleus, but was also seen at what may be nuclear foci (Figure 3.6A). However, these foci do not show any significant co-localisation with KAP-1 in NIH/3T3 cells. In differentiated OS25 cells, Zfp37 was also nuclear diffuse and was not seen to localise to specific foci in experiments carried out by Dr Heidi Sutherland (data not shown). Like Zfp37, Zfp90 also localised to the nucleus in NIH/3T3 cells and gave a fine speckled appearance (Figure 3.6B). There was no significant co-localisation of Zfp90 with KAP-1 in NIH/3T3 cells, and Zfp90 was not seen to localise to KAKA foci in differentiated OS25 cells (data not shown).

An antibody was also received from Dr. Pier Paolo Di Fiore (IFOM, Milan), which was raised against the histone ubiquitin ligase Uhrf1 (Citterio *et al*, 2004). This antibody was raised in rabbit, which was immunised using the full-length protein (Citterio *et al*, 2004). Uhrf1 localised to the nuclei of mouse cells, and in approximately half the cells analysed gave a strong signal from the heterochromatic DAPI bright spots (arrowed in Figure 3.6C). Where the protein was not localised to

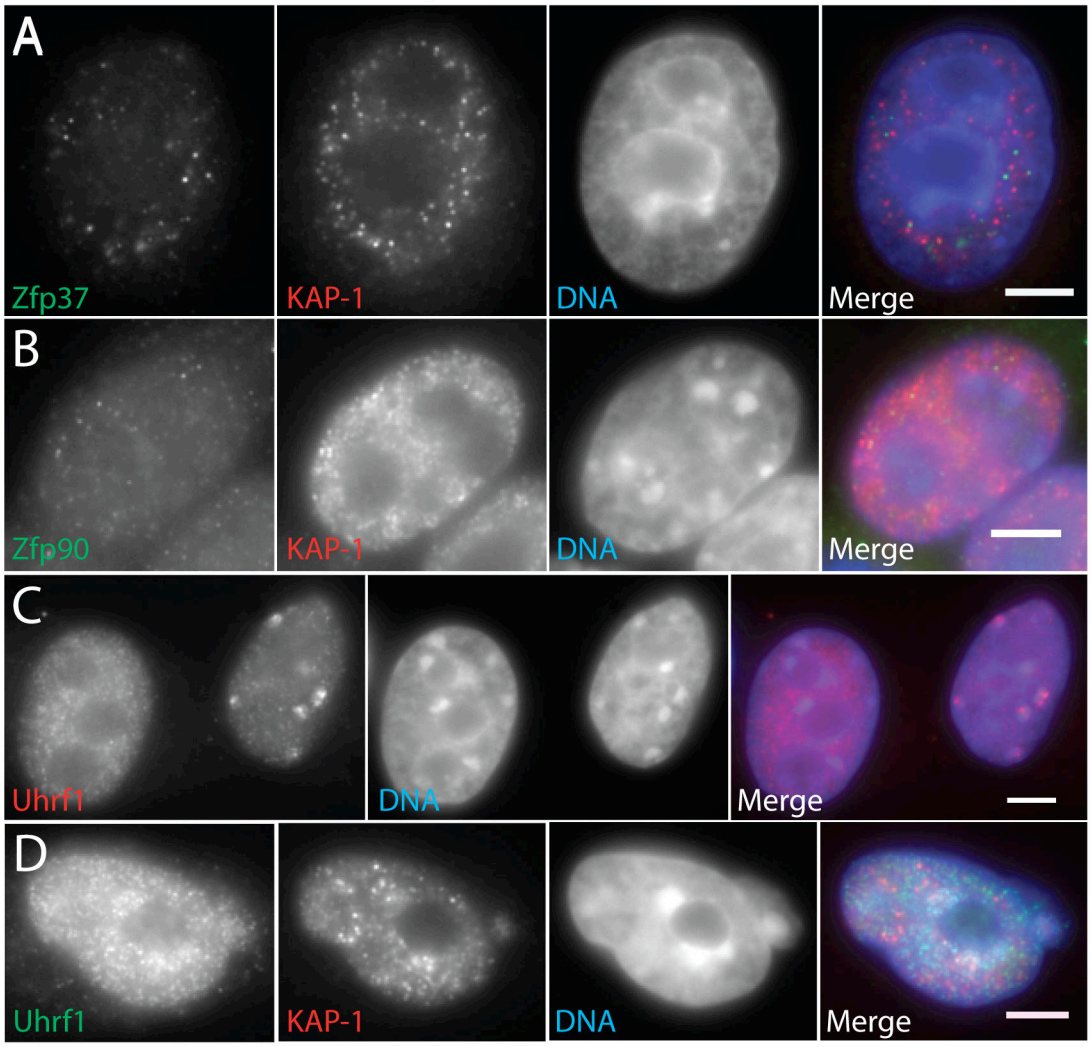


Figure 3.6. Localisation of Zfp37, Zfp90 and Uhrf1 in NIH/3T3 cells. Cells were incubated with A) rabbit α-Zfp37 and mouse α-KAP; B) rabbit α-Zfp90 and mouse α-KAP-1; C) rabbit α-Uhrf1; and D) rabbit α-Uhrf1 and mouse α-KAP-1. Cells in A), B) and D) were incubated with FITC-α-rabbit and Texas red-α-mouse secondary antibodies; cells in C) were incubated with Texas red-α-rabbit secondary antibody. All cells were stained with DAPI to visualise DNA. Scale bars of 5μm are shown.

the heterochromatic bright spots, it was diffuse throughout the nucleoplasm, but did not co-localise with KAP-1 (Figure 3.6D).

These experiments therefore suggested that like in undifferentiated ES cells, Zfp647 is not particularly found in KAKA foci in NIH/3T3 cells, but that these structures are present in pMEFs. The two other KRAB zinc finger proteins tested, Zfp37 and Zfp90, did not localize to KAKA foci in either NIH/3T3 or differentiated ES cells, suggesting that recruitment to the KAKA foci may be specific to specific proteins in certain cell types.

3.4. Rbak Antibody

Given the dearth of good antibodies against mouse KRAB zinc-finger proteins, I decided to generate an antibody against one of those from the yeast two-hybrid screen. As the majority of the _-KRAB zinc finger protein antibodies that have been generated were raised in rabbit, my antibody would be raised in Guinea pig, to allow co-localisation studies in mouse nuclei with rabbit, mouse and sheep antibodies. I chose to raise the antibody against the RB-associated KRAB repressor protein (Rbak), which I found interacting with Zfp647 in the yeast two-hybrid screen (Table 3.1). This 711 amino acid protein shows the typical KRAB zinc-finger protein structure, with an N-terminal KRAB domain and 15 zinc finger motifs (Figure 3.2). The *Rbak* gene is found on mouse chromosome 5, band G2, and is not part of a large multi-gene cluster. One KRAB zinc-finger protein gene, *Zfp12,* is found in the near vicinity of *Rbak*, approximately 60kb away. Zfp12 and Rbak show only 32% peptide sequence identity over their linker regions, although their KRAB and zinc finger domains are more similar.

RBAK was originally identified as an interacting partner of the retinoblastoma gene product, RB, in a yeast two-hybrid screen using human proteins (Skapek *et al*,

2000). This interaction occurs through the linker region of RBAK. It has also been documented that RBAK interacts with the androgen receptor (AR) through its carboxy terminus, again in a yeast two-hybrid screen, backed up by co-immunoprecipitation and mammalian two-hybrid data (Hofman *et al*, 2003). In this case, RBAK appears to function as a coactivator of AR rather than a transcriptional repressor. Skapek *et al* raised an antibody against the human RBAK protein, but did not use this resource to analyse the localisation of the endogenous protein (Skapek *et al*, 2000). Instead, they expressed myc-tagged Rbak and detected the protein using _-myc antibodies. It was reported that myc-Rbak localised to the nucleus, and was possibly excluded from the nucleolus. The linker sequences of the human and mouse Rbak proteins show only 54% identity at the amino acid level, therefore the mouse _-human RBAK antibody is unlikely to detect the murine Rbak protein.

3.4.1. Generation of Rbak Antibody

Primers were designed to amplify the DNA sequence encoding amino acids 69-257 of Rbak, which comprise the linker region between the KRAB and zinc finger domains (Figure 3.7A). The Rbak-linker PCR product was cloned into the pGEX-4T-1 vector (hereafter known as pGEX), N-terminally tagging the protein with GST. Competent BL21 *E. coli* cells were transfected with this vector, and were induced to express GST-Rbak by treatment with IPTG for 3 hours (Section 2.5.1). The protein's solubility was tested by making whole cell, soluble and insoluble extracts from these cells, and running them out by PAGE (Figure 3.7B). pGEX expresses GST protein when induced, seen at 28kDa. The linker region of Rbak has a predicted molecular weight of 21kDa. Coupled with the 28kDa GST protein, the Rbak linker antigen should have a molecular weight of approximately 49kDa. A band corresponding to GST-Rbak linker can be seen at ~50kDa after 3 hours of induction, and the majority of the protein is in the soluble fraction (Figure 3.7B). To produce a sufficient amount of antigen for immunization, a large-scale bacterial culture was induced to express the GST-Rbak linker. This protein was purified from the bacterial lysate by incubation with glutathione beads, to which GST binds, and was then concentrated

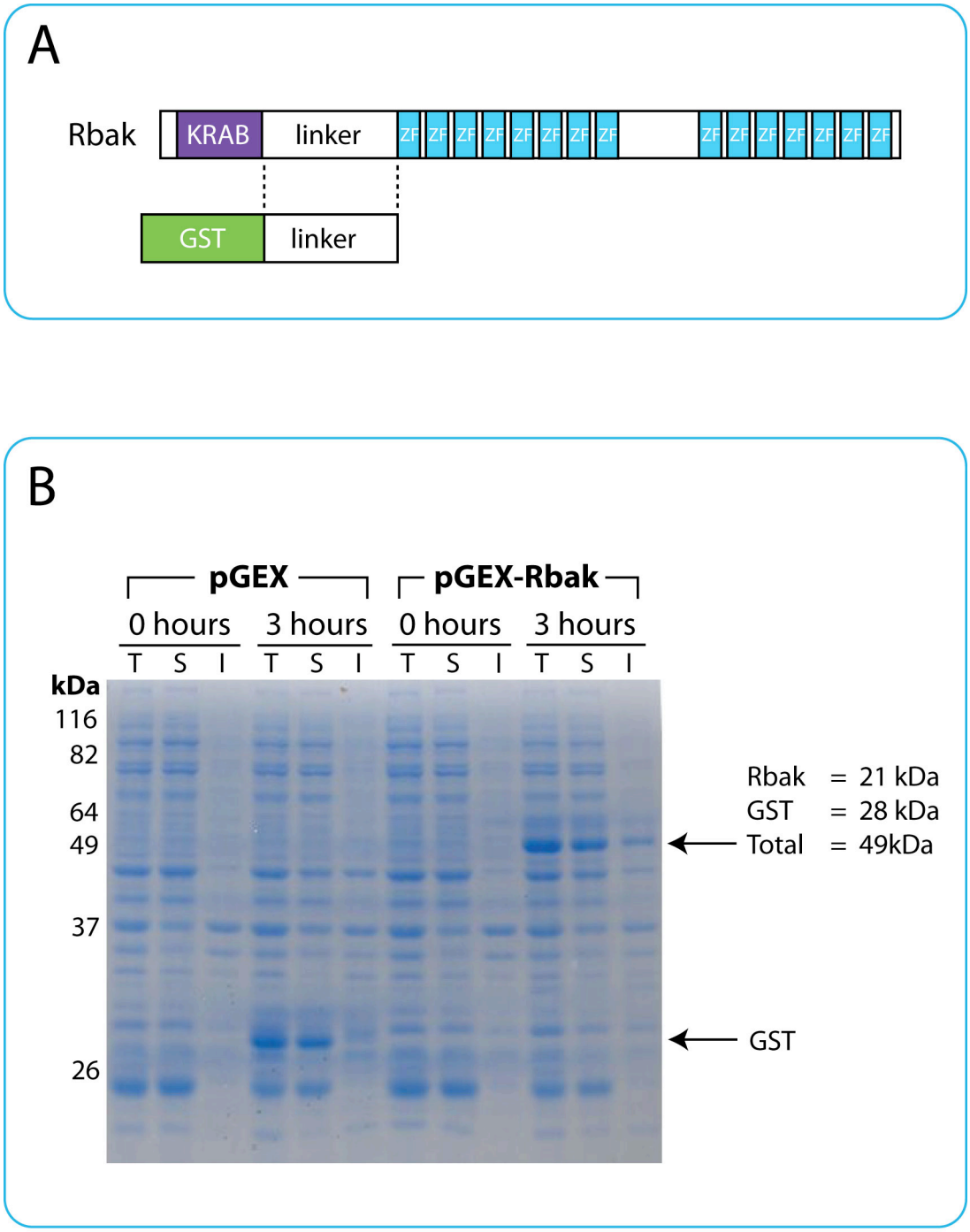


Figure 3.7. Protein structure and expression of the GST-Rbak linker antigen. A) The 188 amino acid linker region was cloned into the pGEX-4T-1 (pGEX) vector, creating an N-terminal GST-fusion protein. B) pGEX and pGEX-Rbak were transformed into *E.coli* cells, and induced with IPTG at 37˚C for 3 hours. Total, soluble and insoluble extracts from uninduced and induced bacteria carrying each plasmid were run out on an SDS-polyacrylamide gel and stained with GelCode blue. T: total cell extract; S: soluble cell extract; I: insoluble cell extract.

and sent to Eurogentec. The antigen was initially used to immunise two Guinea pigs. Each animal was injected four times with the antigen, on days 0, 14, 28 and 56. Four bleeds were taken from the animals: one pre-immune serum, two bleeds from days 38 and 66, and a final bleed from day 87. Unfortunately one Guinea pig died prior to the final bleed being taken, therefore the expression of the antigen and immunisation of one Guinea pig was repeated. The first animal to be immunised was labelled SKC026, the second was SKC047.

To assess whether the final bleed serum contained antibodies specific for Rbak, the pre-immune and final bleed sera were diluted and used in a western blot against the Rbak linker antigen. Whole cell lysates from bacteria carrying the pGEX vector made after 4 hours of induction, and from bacteria transformed with the pGEX-Rbak linker vector (prior to and following induction) were western blotted with the sera diluted at 1:2500. The autoradiographs from the experiments using pre-immune and final bleed sera from animal SKC026 are shown in Figure 3.8A, and those from animal SKC047 are shown in Figure 3.9A. Neither pre-immune serum detected any proteins in the bacterial lysates. Both final bleed sera strongly detect a band of approximately 50kDa (Figures 3.8A and 3.9A), which corresponds to GST-Rbak linker. As the animals were immunized with a GST-fusion protein, it was likely that antibodies would be raised against GST as well as the Rbak portion of the antigen: these antibodies are clearly present in animal SKC026's final bleed serum, seen in the pGEX 4 hour lane (Figure 3.8A). Antibodies present in this serum also detect a non-specific band at 80kDa (Figure 3.8A).

The final bleed sera were then tested against mouse cell extracts, to see whether they could bind to endogenous Rbak. The predicted size of the full-length protein is 82kDa, and Skapek *et al* found that their antibody against human Rbak detected a band at 80kDa (Skapek *et al*, 2000). NIH/3T3 whole cell, nuclear and cytoplasmic extracts were probed with the pre-immune and final bleed sera from both Guinea pigs, again at a range of dilutions (1:1000 – 1:5000) (Figures 3.8B and 3.9B). The pre-immune sera from both animals show a minimal amount of signal on the western blots, demonstrating the lack of antibodies capable of binding to mouse proteins in

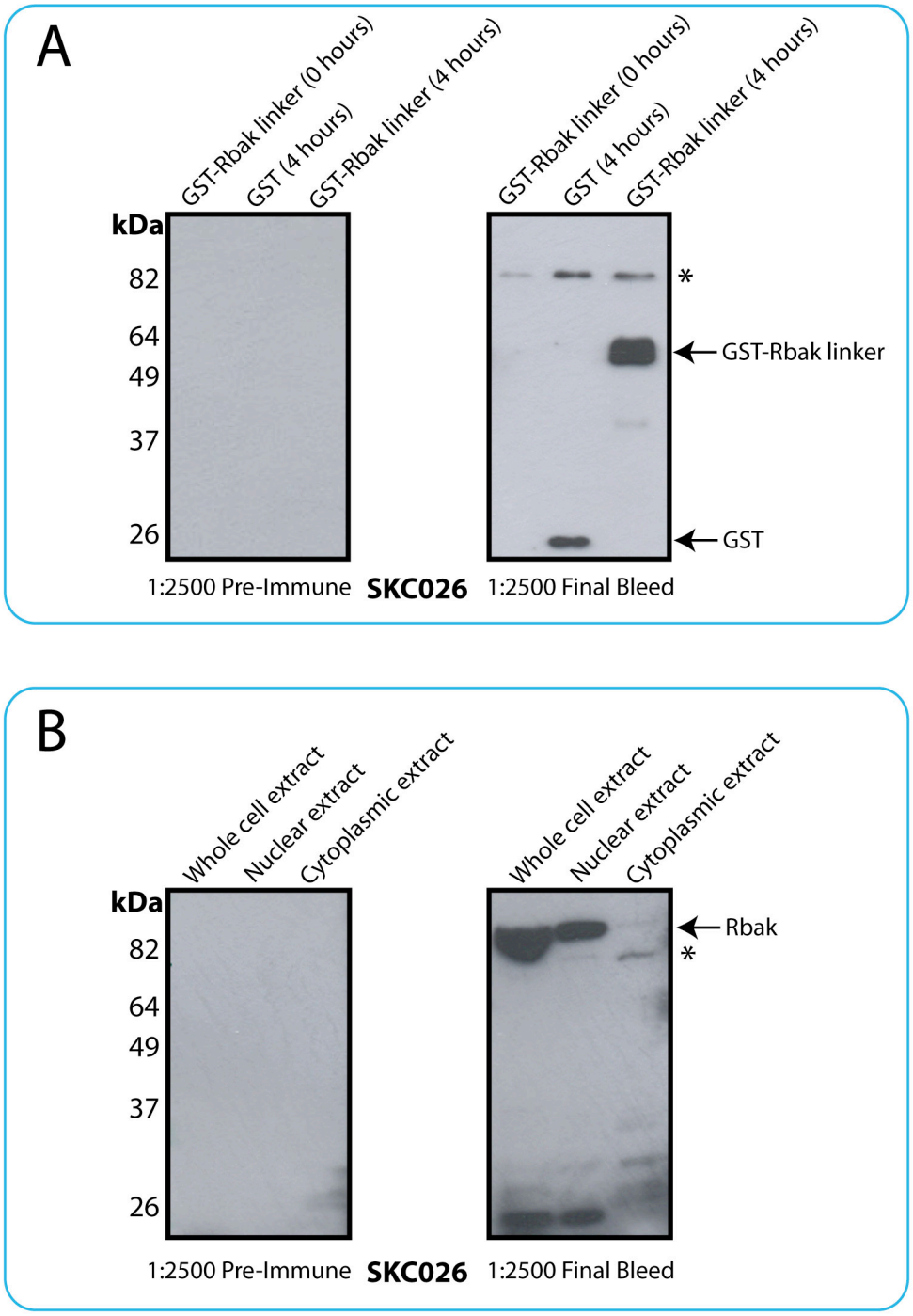


Figure 3.8. Testing of α-Rbak serum from Guinea pig SKC026. A) Western blots of extracts from *E. coli* carrying either the empty pGEX vector (expressing GST) or the pGEX-Rbak linker vector (expressing GST-Rbak linker) probed with Guinea pig pre-immune or final bleed serum. Extracts were made from both uninduced (0 hours) and induced (4 hours) bacterial cultures; only the induced sample is shown for cells carrying pGEX. Arrows indicate GST and GST-Rbak; asterisk indicates an unknown unspecific band detected by the antibody. B) Western blots of NIH/3T3 whole cell, nuclear and cytoplasmic extracts probed with a 1:2500 dilution of either Guinea pig pre-immune or final bleed serum. Arrow indicates presumed Rbak band detected by the antibody; asterisk indicates an unknown cytoplasmic protein.

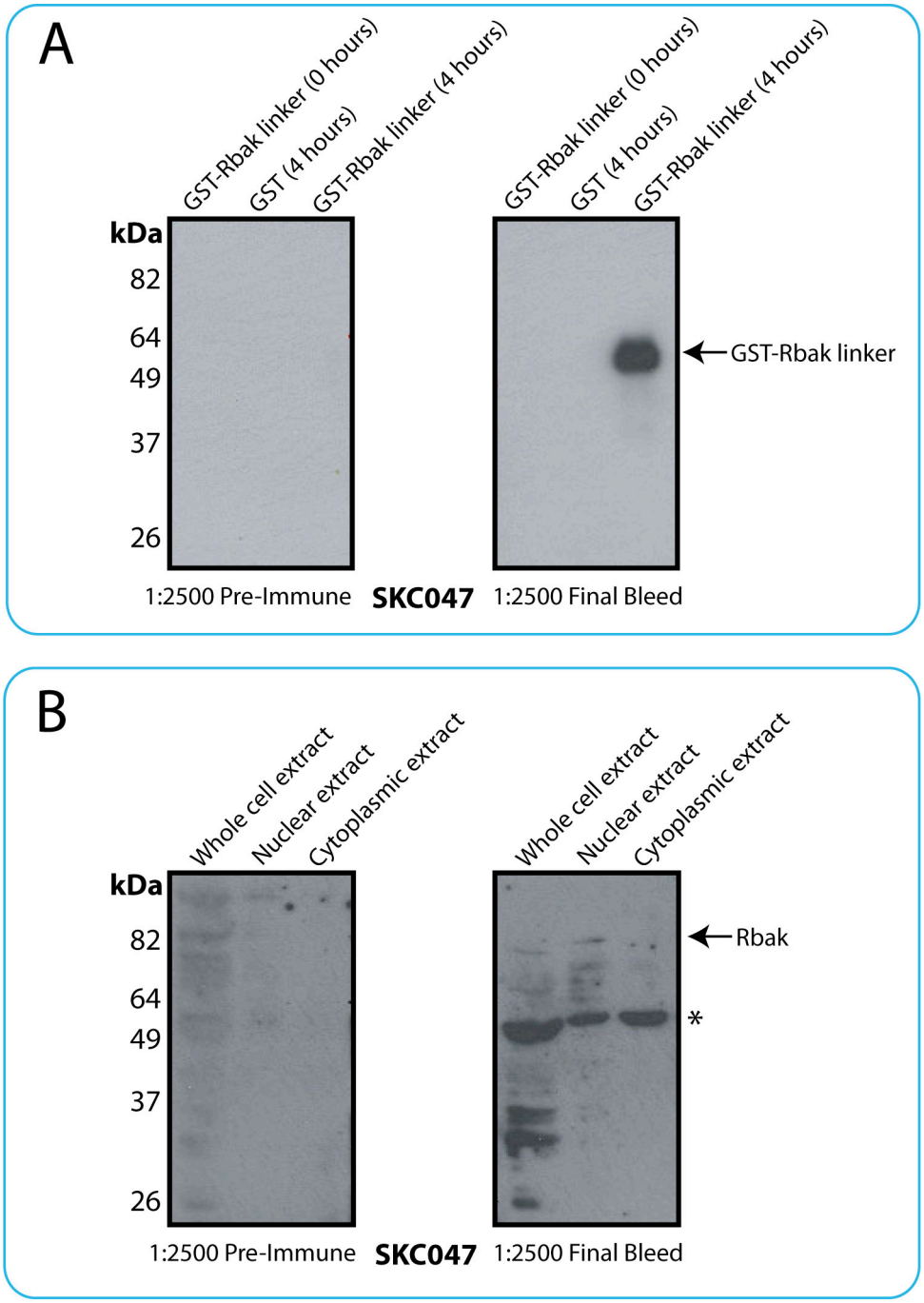


Figure 3.9. Testing of α-Rbak serum from Guinea pig SKC047. A) Western blots of extracts from *E. coli* carrying either the empty pGEX vector (expressing GST) or the pGEX-Rbak vector (expressing GST-linker) probed with Guinea pig pre-immune or final bleed serum. Extracts were made from both uninduced (0 hours) and induced (4 hours) bacterial cultures; only the induced sample is shown for cells carrying pGEX. Arrow indicates GST-Rbak. B) Western blots of NIH/3T3 whole cell, nuclear and cytoplasmic extracts probed with a 1:2500 dilution of either Guinea pig pre-immune or final bleed serum. Arrow indicates the predicted size of endogenous Rbak; asterisk indicates an unknown protein detected by final bleed serum.

these animals (Figures 3.8B and 3.9B) The final bleed serum from Guinea pig SKC026 strongly detects a band at approximately 90kDa in the whole cell and nuclear extracts (Figure 3.8B), which would indicate that this serum contains antibodies that can detect the endogenous mouse Rbak protein. The final bleed serum from Guinea pig SKC047 also seems to detect this band, but at a much lower level compared to the SKC026 serum (Figure 3.9B). The SKC047 serum detects another band at 50kDa more strongly; this unknown protein is found in both the nuclear and cytoplasmic fractions of the cell (Figure 3.9B). The SKC026 serum also detects a predominantly cytoplasmic band, at approximately 85kDa (Figure 3.10, indicated with asterisk). Because of the presence of a strong nuclear band at the correct size on the western blots carried out with the SKC026 serum, I decided to carry out all further work using this antibody.

3.4.2. Purification of Guinea Pig _-Rbak Serum

The final bleed serum of Guinea pig SKC026 was affinity-purified on an affinity column containing the Rbak linker isolated on sepharose beads (Section 2.5.4). Ideally, the Rbak protein would have been a His_6-tagged construct. This would mean that the antibodies against GST that are present in the serum would not bind to the column, purifying only the antibodies specific to Rbak. A His_6-tagged Rbak linker construct was cloned in the pET32a vector, and transfected into *E. coli* cells. However, when expressed, the His_6-Rbak linker protein was insoluble, and thus could not be bound to sepharose beads to make a column (data not shown). To remove antibodies against GST, the serum was first passed over an affinity column for another GST fusion protein, GST-p52, which is a splice variant of the lens epithelium-derived growth factor (LEDGF), a protein unrelated to Rbak. The _-Rbak serum could then be purified over a GST-Rbak affinity column.

Purified _-RBAK antibody was first tested by western blot against mouse NIH/3T3 whole cell, nuclear and cytoplasmic extracts in dilutions from 1:200 to 1:1000 (Figure 3.10A show the results using a 1:200 dilution). The western blot of

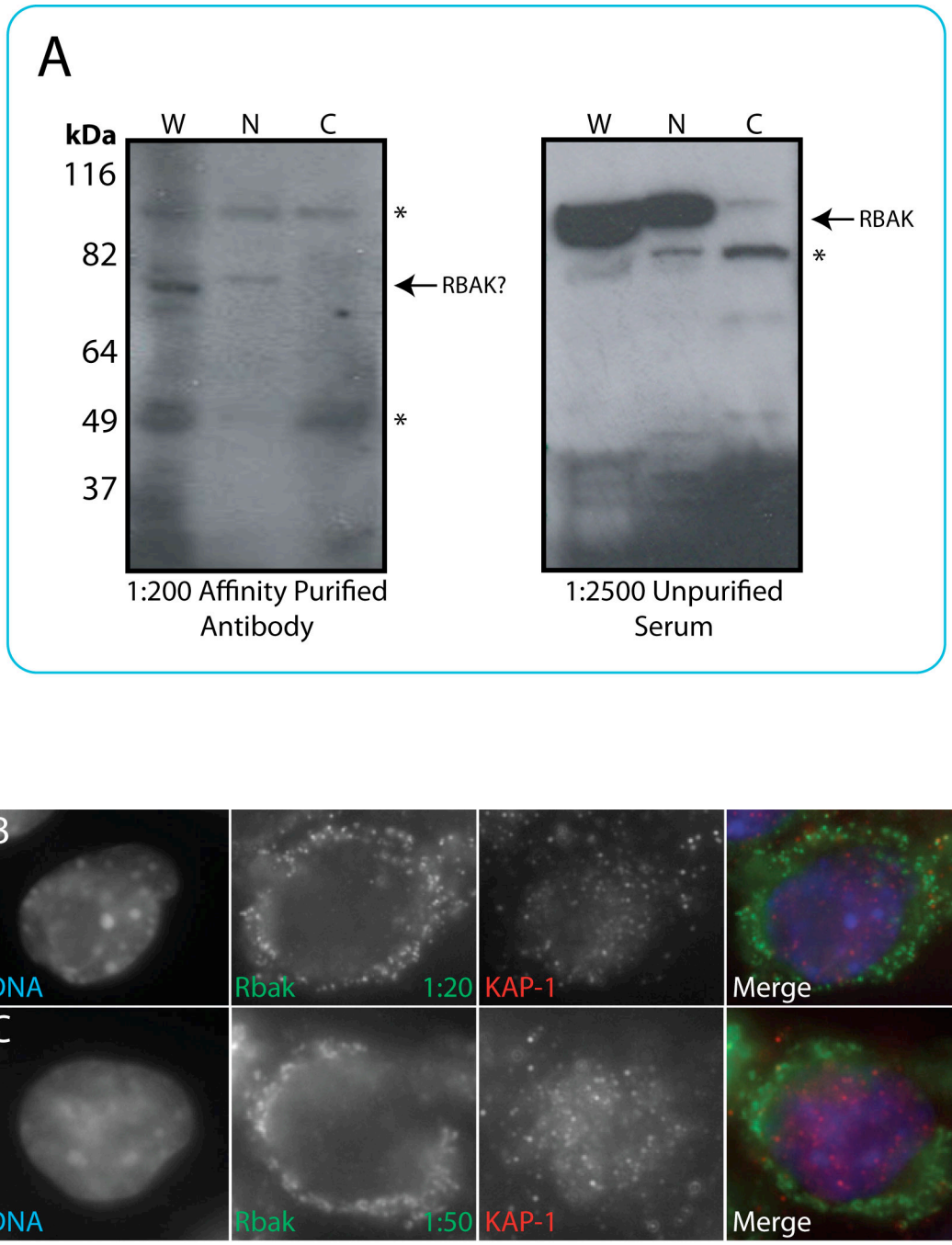


Figure 3.10. Testing affinity purified SKC026 Guinea pig α-Rbak. A) Western blots of whole cell (W), nuclear (N) and cytoplasmic (C) extracts made from NIH/3T3 cells probed with a 1:200 dilution of affinity purified Guinea pig α-Rbak (left panel) or 1:2500 unpurified Guinea pig final bleed serum. Filled arrow indicates possible Rbak band, asterisks indicate non-specific bands. B) and C) Immunofluorescence using affinity-purified antibody from Guinea pig SKC026 final-bleed serum. NIH/3T3 cells were incubated with affinity-purified Guinea pig α-Rbak diluted at B) 1:20 or C) 1:50 in blocking solution, and staining was detected with a FITC-conjugated α-Guinea pig secondary antibody. Cells were also incubated with rabbit α-KAP-1, detected with Texas red-α-rabbit secondary antibody.

the same extracts probed by the unpurifed SKC026 serum is shown alongside these results. By direct comparison, it is clear that the affinity-purified antibody does not give as strong a signal as the unpurified serum. There appears to be a band present at approximately 80kDa in the whole cell and nuclear extracts when blotted with the purified antibody, which may correspond to Rbak. This band is slightly higher in the nuclear extract than in the whole cell extract, which is also the case with extracts probed with the unpurified serum. Non-specific bands are also seen at ~50kDa and 100kDa (Figure 3.10A).

The purified antibody was also checked for its ability to detect Rbak within mouse cells by immunofluorescence. NIH/3T3 cells were incubated with 1:20 – 1:100 dilutions of the purified Rbak antibody and 1:1000 rabbit _-KAP-1 antibody, which were detected with FITC-conjugated _-Guinea pig and Texas red-_-rabbit secondary antibodies. Images of these cells are shown in Figure 3.10B and C. The speckled nuclear pattern of KAP-1 can be seen in both cells shown (Figure 3.10B and C). However, like the immunofluorescence with the unpurified Guinea pig SKC026 (data not shown), the staining seen with the affinity-purified antibody is almost completely cytoplasmic, and there is hardly any signal from the nucleus. The detected protein cannot be the 90kDa band seen in NIH/3T3 extracts, as that protein was almost completely nuclear (Figure 3.10, right-hand panel). There is no problem with the permeability of the cells, as the _-KAP-1 antibody has stained the nucleus, therefore it is unlikely that the SKC026 antibody has been prevented from entering the nucleus. Thus either the serum is staining an unknown cytoplasmic protein in this immunofluorescence experiment, or Rbak is not the nuclear protein observed by western blot, but is instead the lower, more cytoplasmic band. Although we predict that Rbak is nuclear, and tagged RBAK been observed in the nucleus (Skapek *et al*, 2000), it may be that mouse Rbak localises to the cytoplasm. Because I was unsure which of these hypotheses was correct, I did not carry out any further immunofluorescence experiments with this antibody.

3.5. Co-Immunoprecipitation of Endogenous Zfp647 Interactors

My yeast two-hybrid screen has revealed many putative Zfp647 interactors. However, these proteins were identified using a non-endogenous system, thus they may not interact with Zfp647 *in vivo*. I therefore attempted to identify Zfp647-interacting proteins in an endogenous manner. Because antibodies specific to Zfp647 had been successfully generated in the lab, I chose to immunoprecipitate Zfp647, and try to identify putative interacting proteins by mass spectrometry. If this method generated any results that overlapped with those from the yeast two-hybrid screen, it would serve as an independent verification of these results. The co-immunoprecipitations were carried out in NIH/3T3 cells. Previous work has shown that Zfp647 is not recruited to KAKA foci in this cell line, as it is in pMEF or differentiated ES cells (Figure 3.5). However, carrying out a large-scale immunoprecipitation was not viable with differentiated ES cells, given the small amount of protein gained from a flask of these cells, and the large amount of material required for the mass spectrometry. Similarly, I experienced difficulties growing pMEF cells for high numbers of passages, which was also needed to generate enough material for both the immunoprecipitation and the control. Therefore I decided to carry out the co-immunoprecipitation experiment with NIH/3T3 cells. While this might not tell me proteins that co-localised with Zfp647 at the KAKA foci, it may enable me to gain an insight into possible functions of Zfp647 by revealing putative endogenous interacting proteins, and also other factors that may function in KRAB-mediated transcriptional repression.

3.5.1. Testing of Rabbit and Sheep -Zfp647 Antibodies

I initially tested both the rabbit and sheep antibodies raised against Zfp647 by western blot to show that both could bind to Zfp647, and were not specific to any other proteins in the mouse proteome. Both antibodies were raised against the linker region of Zfp647, and were affinity purified from the final bleed sera using the same region of the protein. Mouse lines have been established by Drs Heidi Sutherland and

Stephanie Briers from the gene-trapped Zfp647 ES cells, allowing the generation of wild type mice and mice both heterozygous and homozygous for gene-trapped Zfp647. I tested the antibodies against protein extracts made from wild-type and homozygous gene-trapped mouse embryos (Figure 3.11). The sheep antibody detects a band at 60kDa, corresponding to the predicted size of Zfp647, in the wild-type protein extract, which is not seen in the gene-trapped mouse extract. Similarly, this band is only detected by the rabbit antibody in the wild-type extract and not the gene-trapped extract. Both the sheep and rabbit antibodies detect this band in a NIH/3T3 cell nuclear extract. The rabbit antibody additionally gives a signal at approximately 40kDa in the mouse gene-trapped extract. This band does not disappear when endogenous Zfp647 is lost, therefore it seems unlikely that it is a degradation or cleavage product of Zfp647. Analysis of whole cell, cytoplasmic and nuclear NIH/3T3 extracts show that the band is cytoplasmic (Figure 3.11), and therefore should not interfere with immunofluorescence studies of the endogenous protein, nor with studies carried out using NIH/3T3 nuclear extracts.

3.5.2. Immunoprecipitation of Zfp647

To successfully co-immunoprecipitate Zfp647-interacting proteins, I first needed to verify that I could immunoprecipitate Zfp647 itself. As the two _-Zfp647 antibodies were raised in different species, this allowed me to immunoprecipitate Zfp647 with one antibody, and test for its presence with the other. Because rabbit antibodies have a much higher affinity to protein G than sheep antibodies have to either protein A or G (Harlow and Lane, 1988), I bound this antibody to protein G sepharose beads, incubated these with NIH/3T3 nuclear extract, and washed the beads (Section 2.4.5). The immunoprecipitated material was then run out by PAGE and western blotted with sheep _-Zfp647 (Figure 3.12). If the immunoprecipitations are prepared for PAGE with reducing agent in the Laemmli buffer, the IgG bands run at approximately the same size as Zfp647, therefore these immunoprecipitations were prepared in the absence of reducing agent. The IgG bands can be seen as a smear

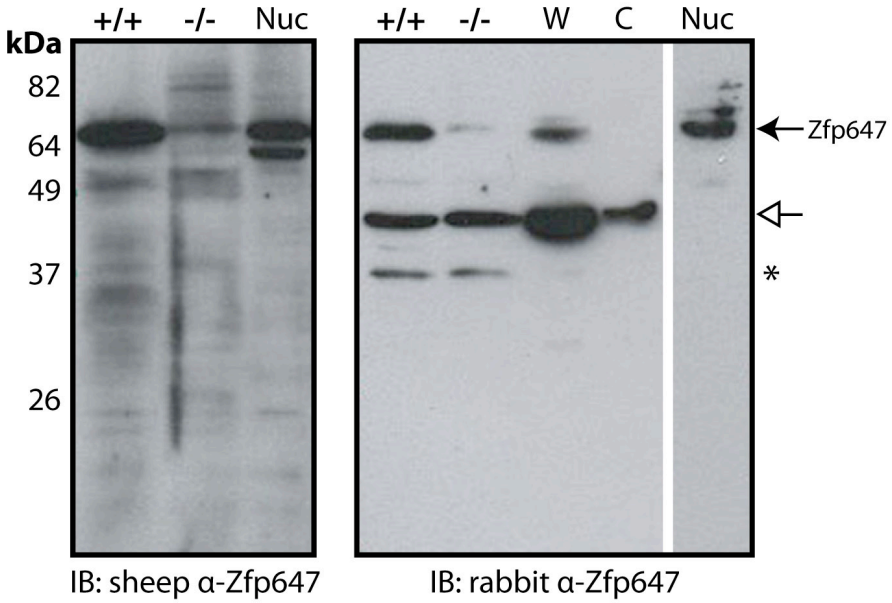


Figure 3.11. Testing of sheep and rabbit α-Zfp647 antibodies. Western blot of protein extracts made from wild-type (+/+) and Zfp647 gene-trap homozygous (-/-) 12.5dpc mouse embryos blotted with both sheep and rabbit α-Zfp647 antibodies. Both antibodies were also tested against NIH/3T3 nuclear extracts (Nuc), and the rabbit α-Zfp647 antibody was tested against NIH/3T3 whole cell (W) and cytoplasmic (C) extracts. Arrow indicates Zfp647; unfilled arrow indicates a cytoplasmic band also detected by the rabbit antibody; asterisk indicates a non-specific band present in embryonic protein extracts.

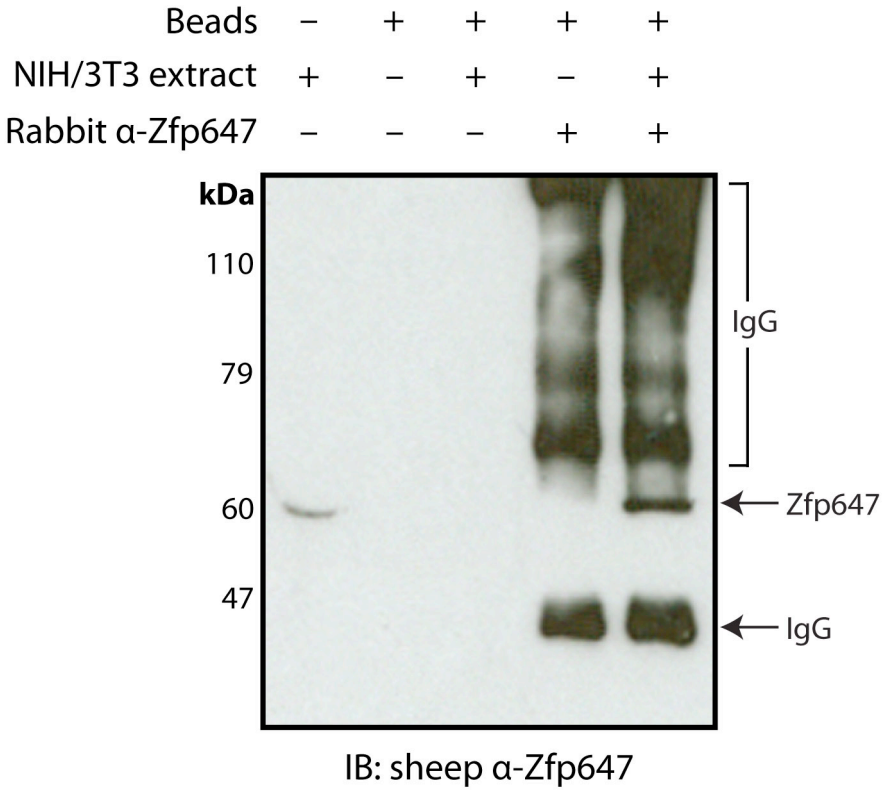


Figure 3.12. Immunoprecipitation of Zfp647. Western blot of immunoprecipitation carried out with rabbit α-Zfp647 antibody, blotted with sheep α-Zfp647. Zfp647 and IgG bands are indicated.

above and below the predicted size of Zfp647. The sheep antibody recognizes the 60kDa Zfp647 band in the input, as shown previously (Figure 3.11). The same sized band can clearly be seen between the IgG bands only in the immunoprecipitation lane, therefore I have shown that this antibody can successfully immunoprecipitate Zfp647 from nuclear extracts. I also western blotted the immunoprecipitated material with mouse _-KAP-1 antibody, to see whether a known protein interactor could be co-immunoprecipited by this antibody. However, despite numerous attempts at this experiment, I was unable to confirm the presence of KAP-1 in the immunoprecipitated material (data not shown). This may be because Zfp647 does not bind to KAP-1 in NIH/3T3 cells, or the antibody may interfere with the interaction between the two proteins, therefore preventing the co-immunoprecipitation of Zfp647-KAP-1 complexes. Alternatively, the Zfp647 epitope may be masked by the KRAB-KAP-1 interaction, given the proximity of the linker region to the KRAB domain.

3.5.3. Mass Spectrometry of Co-Immunoprecipitated Proteins

To identify potential interacting proteins, the immunoprecipitations were run out by PAGE and silver stained. Any bands present in the immunoprecipitated material that were not in the control lanes could be potential interactors, and would be cut from the gel and subjected to mass spectrometry. Although leaving out reducing agent from the SDS loading buffer when the immunoprecipitations were run out by PAGE allowed the clear identification of Zfp647 by western blot, I included reducing agent in the immunoprecipitations that were run out and silver stained to remove the smear from the lanes. I also scaled up the immunoprecipitations by increasing the amount of input, beads and antibody used, in order to maximize the signal of any co-immunoprecipitated material. In total, cells confluent in 8x 14cm tissue culture plates were harvested for each immunoprecipitation. Two immunoprecipitations were carried out simultaneously, in the presence and absence of the chemical N-ethylmaleimide (NEM). NEM is a cysteine protease inhibitor, and preserves ubiquitin-like modifications on proteins, a topic that will be further discussed in

Chapter 4. Both immunoprecipitations and their controls were run out by PAGE on a pre-cast gel and silver stained (Figure 3.13). The control lanes lacking antibody show a high degree of background: most of the bands present in the immunoprecipitation lanes are also present in the control lanes (compare lanes 4 and 5 with lanes 2 and 3). However, a number of bands were present only in the immunoprecipitated lanes, and three of these (marked in Figure 3.13) were excised from the gel and sent to the SIRCAMS mass spectrometry service at the University of Edinburgh Chemistry Department. Here, the samples were digested with trypsin and subjected to matrix-assisted laser desorption/ionization time of flight (MALDI-TOF) mass spectromeric analysis, whereby the tryptic fragments are ionised by a laser, and their molecular weight is calculated by their time of flight through a chamber of known length. From piecing together different peptide fragments, the computer program ProFound (http://prowl.rockefeller.edu/) can generate 'hits': proteins that contain peptides of those sizes. The results received for each band are shown in Table 3.8. The expectation score represents how likely it is that this protein is a true match: an expectation value of 1 means that you would expect one match if you used the mass spectrometry data for a particular protein to search a database that you knew did not contain that protein. The smaller the expectation number, the greater the likelihood that the match is true: an expectation value of 0.01 means that you would expect one match to your data in 100 databases that did not contain your protein. Therefore the expectation value functions similarly to a p-value. The % value shows the percentage of the protein covered by the recovered peptides; pI shows the estimated protein pI value; and kDa represents the protein's estimated molecular weight in kDa.

Using the typical value of statistical significance of 0.05, none of the protein results obtained are statistically significant. One result almost reaches statistical significance, and this is the top-ranked result for Band 3, Jarid1b. Jarid1b is a jumonji domain-containing protein, also known as PLU-1. The coverage of the protein by the recovered peptides was not high, at 13% (a good hit is considered to be one that gives 20% coverage), but the peptides cover a wide range over the length of the protein, shown in Figure 3.14. At 179kDa, the protein is also approximately the correct size for the band on the gel, unlike other results in Table 3.8 (for example,

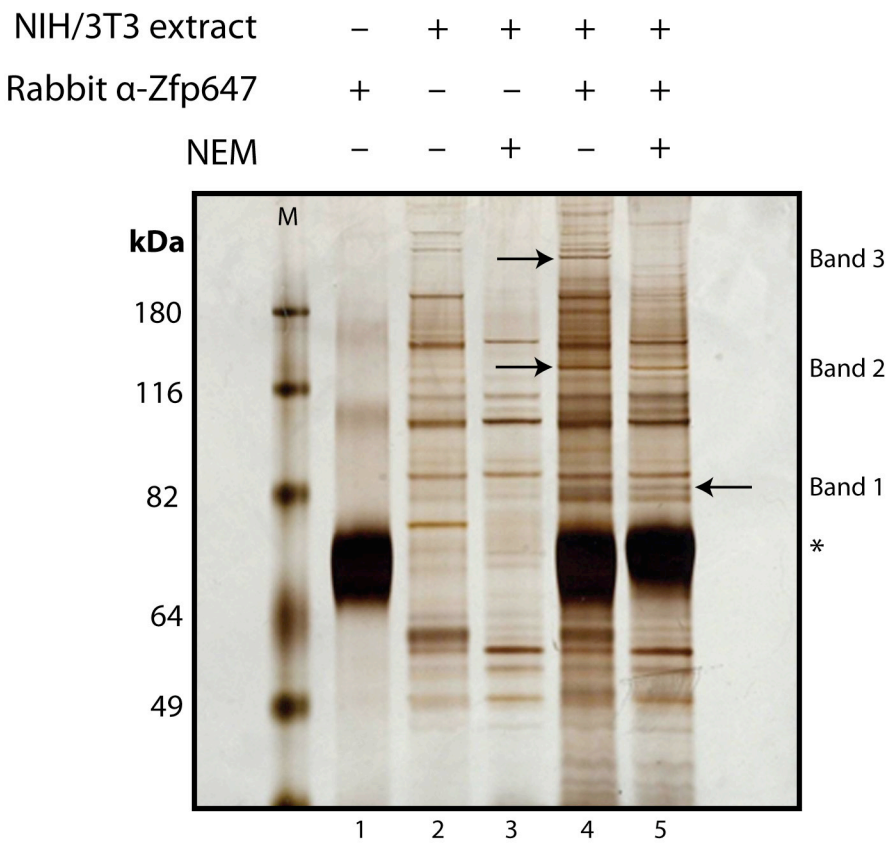
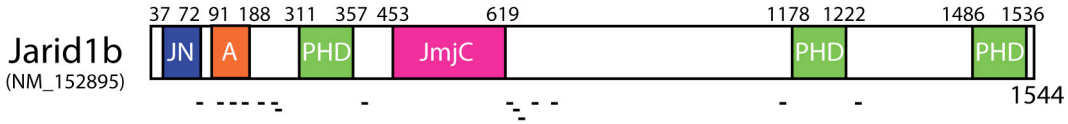


Figure 3.13. Immunoprecipitation of Zfp647 for mass spectrometry. Silver-stained gel of Zfp647 immunoprecipitations carried out in the presence or absence of N-ethylmaleimide (NEM). Bands excised for mass spectromeric analysis are indicated with arrows; asterisk indicates IgG bands. M indicates marker lane.

Band	Rank	Protein (Accession No.)	Expectation	%	pI	kDa
1	1	Protein tyrosin kinase 2 beta (Ptk2b) (NM_172498)	0.39	13	5.7	117
1	2	Coiled-coil domain containing 18 (Ccdc18) (NM_028481)	0.72	13	5.2	48
1	3	Protein tyrosine phosphatase, receptor type, f polypeptide interacting protein alpha 1 (Ppfia1) (NM_001033319)	0.72	13	5.7	117
1	4	Deleted in liver cancer 1 (Dlc1) (NM_001127446)	0.83	11	7.5	124
2	1	Phosphodiesterase 10A (Pde10a) (NM_011866)	0.20	22	6.3	90
2	2	Deleted in liver cancer 1 (Dlc1) (NM_001127446)	0.35	15	7.5	124
2	3	Protein tyrosine phosphatase, receptor type, T (Ptprt) (NM_021464)	0.84	7	6.4	164
2	4	Junction plakoglobin (Jup) (NM_010593)	0.88	16	6.0	69
3	1	Jumonji, AT rich interactive domain 1B (Jarid1b) (NM_152895)	0.070	13	6.0	179
3	2	Protein tyrosine phosphatase, non-receptor type 13 (Ptpn13) (NM_011204)	0.23	13	8.5	116
3	3	MutS homolog 6 (Msh6) (NM_010830)	0.48	12	6.3	153
3	4	Myosin, heavy polypeptide 9, non-muscle (Myh9) (NM_022410)	0.53	11	5.4	156
3	5	Deleted in liver cancer 1 (Dlc1) (NM_01127446)	0.62	13	7.5	125

Table 3.8 Mass spectrometry results for Zfp647 co-immunoprecipitation. Table shows proteins identified as hits in mass spectrometry of bands shown in Figure 3.14. % indicates the proportion of the protein covered by the identified peptides.



64-86	VRPPPDWWQPPFACDVDKLHFTPR
120-126	IPHVERK
136-150	LVAEEGGFAVVCKDR
159-172	MGFAPGKAVGSHIR
207-216	EYKPHDIPQR
236-252	AEAMNIKIEPEEATEAR
243-257	IEPEEATEARTHNLR
376-390	DYTLRTFGEMADAFK
624-635	YCVFSHDEMICK
640-663	ADVLDVVVASTVQKDMAIMIEDEK
654-666	DMAIMIEDEKALR
680-694	MDFELLPDDERQCIK
721-733	ELCSCPPYKYNLR
1152-1166	EMEALQSLRFANEEK
1246-1253	LPEGDALR

Figure 3.14. Structure of Jarid1b. Upper figure shows domains contained within Jarid1b, and the location of the tryptic peptides. Lower figure shows peptide sequences and their location within the Jarid1b sequence. JN: JmjN domain; A: AT-rich DNA binding (ARID) domain; PHD: plant homeodomain.

Ptpn13). Proteins belonging to the Jarid1 family are H3K4 demethylases (Christensen *et al*, 2007; Iwase *et al*, 2007; Yamane *et al*, 2007), and can therefore function as transcriptional repressors. Human JARID1B possesses H3K4 demethylase activity (Yamane *et al*, 2007), and is upregulated in breast and prostate cancer (Lu *et al*, 1999; Xiang *et al*, 2007). While this result is not statistically significant, it would be of worth to try to verify whether Zfp647, KAP-1 or another protein involved in KRAB zinc finger protein-mediated transcriptional repression interacts with Jarid1b, or other jumonji domain protein, as I have identified two jumonji domain proteins as possible Zfp647-interactors (Section 3.2.5).

3.6. Co-Immunoprecipitation of Tagged Zfp647

Despite my yeast two-hybrid screen yielding results that were probably 'true' interactions, none of these results were verified by co-immunoprecipitating endogenous Zfp647 interactors from NIH/3T3 cells. I therefore decided to try and validate the yeast two-hybrid results using a separate method. The first interaction I focussed on was that of the self-interaction of Zfp647. Although it had already been found that the endogenous protein behaves differently to an overexpressed Zfp647 fusion construct with respect to its localisation, I chose to look at the interaction between two fusion proteins as I did not believe that N-terminal protein tags would interfere with the interaction domain in Zfp647's C-terminus. Analysing the interaction between two fusion proteins would also be the simplest way to detect the homo-oligomerisation of Zfp647.

3.6.1. Expression of Tagged Zfp647 Constructs

I initially planned to clone and express two soluble, tagged forms of Zfp647 either by bacterial expression, or in mouse tissue culture cells. This would allow me

to immunoprecipitate one of the proteins with an antibody against one of the tags, and detect for an interaction with the alternatively tagged Zfp647 protein by immunoblotting the immunoprecipitated material with an antibody against the other tag. I attempted to express three different forms of Zfp647, with GST, T7 and His_6 tags.

3.6.1.1. Cloning and Expression of Bacterial GST-Zfp647

To show that ZBRK1 homo-tetramerises, Tan *et al* expressed ZBRK1 fused to maltose binding protein (MBP) in *E. coli* (Tan *et al*, 2004b). This protein was bound to amylase resin, incubated with *in vitro*-expressed, radio-labelled ZBRK1, and their interaction was demonstrated by the detection of a radioactive signal from the resin. Similarly, Nielsen *et al* immobilised GST-NSD1, a histone methyltransferase, on glutathione-sepharose, and confirmed its ability to bind to the KRAB zinc-finger protein Nizp1 by incubating it with bacterially-expressed His_6-Nizp1 and blotting the resin-bound proteins with an _-His_6 antibody (Nielsen *et al*, 2004). I chose to express Zfp647 with a GST tag, immobilise this protein on glutathione beads, and carry out the immunoprecipitation in a similar manner using, as a prey protein, *in vitro*-expressed Zfp647.

The full length sequence of Zfp647 was cloned into the pGEX vector to give the protein an N-terminal GST tag. This plasmid was transfected into BL21 *E. coli* cells, which were induced to express the protein at 37°C for 4 hours. To carry out a co-immunoprecipitation experiment, the protein must be soluble, as any insoluble matter is removed from the lysate prior to incubation with beads. Whole cell and soluble extracts were therefore prepared from these bacterial cultures before and after induction, and analysed by PAGE (Figure 3.15A). A culture of bacteria carrying the empty pGEX vector was also induced as a control, and soluble GST protein is clearly expressed by these cells 4 hours after induction with IPTG. GST-Zfp647 of the expected size (80kDa) is also expressed following induction of the culture, and can be seen in the total cell extract. However, unlike the GST protein, GST-Zfp647 is not

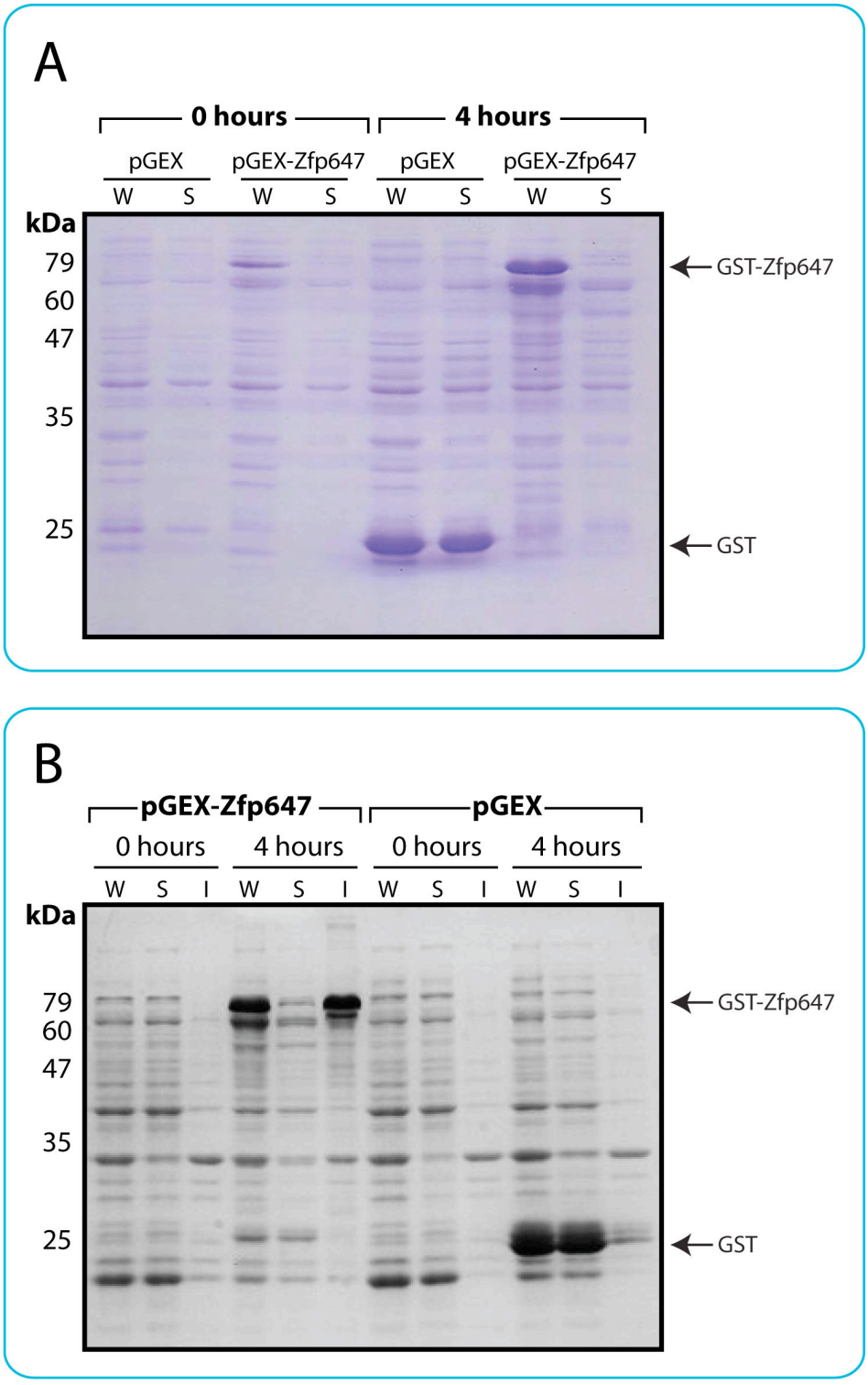


Figure 3.15. Bacterial expression of GST-tagged Zfp647. A) Coomassie-stained PAGE gel of lysates of bacteria expressing the pGEX-4T-1 (pGEX) and pGEX-Zfp647 vectors at 37°C. Bacteria were induced by treatment with IPTG for 4 hours. Whole cell (W) and soluble (S) extracts made from cells prior to and following induction are shown. Arrows indicate GST and GST-Zfp647. B) Coomassie-stained PAGE gel of lysates of bacteria expressing pGEX and pGEX-Zfp647 at 30°C. Bacteria were induced with IPTG for 4 hours. Whole cell (W), soluble (S) and insoluble (I) extracts made from cells prior to and following induction are shown.

present in the soluble fraction when expressed at 37°C, therefore this protein is insoluble under these conditions. This may be due to the incorrect folding of the protein resulting from high levels of expression, which can lead to the formation of insoluble protein aggregates. The insoluble protein lysates were therefore treated with up to 5M urea, which denatures the non-covalent bonds of the protein's structure, and should help to remove protein aggregates. The proteins were then allowed to re-fold, ideally solubilising Zfp647. However, when analysed, the urea-treated lysates did not yield any soluble GST-Zfp647 protein (data not shown).

GST-Zfp647 was then expressed at a range of temperatures, which may slow down the level of protein production, promoting the correct folding of Zfp647 and increasing its solubility. Expression of GST-Zfp647 was attempted at 30°C, 23°C and 16°C. Protein was only successfully harvested from bacteria growing at 30°C and 23°C; at 16°C, the culture did not grow well, barely completing one doubling within 3 hours. Figure 3.15B shows the total, soluble and insoluble extracts from cultures grown at 30°C; the results from the culture grown at 23°C were identical. GST-Zfp647 is clearly expressed, but is insoluble. As bacterially-expressed Zfp647 is not soluble at 37°C, 30°C or 23°C, full-length GST-Zfp647 could not be used in a co-immunoprecipitation experiment.

3.6.1.2. Cloning and Expression of Mammalian His_6-Zfp647 and T7-Zfp647

An alternative to producing Zfp647 bacterially was to clone the protein into vectors that would allow expression in mammalian cells. I created two such Zfp647 constructs: one in the vector pCGT7, giving the protein an N-terminal T7 tag (Cáceres *et al*, 1997); and one in the vector pcDNA3.1(+) (hereafter referred to as pcDNA). To create the insert for the latter plasmid, the full-length Zfp647 sequence was amplified by PCR using a 5' primer containing the nucleotide sequence CAT CAT CAC CAT CAC CAT, tagging the protein with six histidine residues. I planned to co-transfect the two vectors into NIH/3T3 cells, make protein extract from these cells and immunoprecipitate Zfp647 with an _-T7 antibody. The immunoprecipitated

material would then be western blotted with an _-His_6 antibody to detect the ability of the protein to oligomerise. First, the two plasmids were independently transfected into NIH/3T3 cells, and soluble and insoluble nuclear extracts were made, in order to ascertain the solubility of the tagged proteins.

Extracts were made at a range of NaCl concentrations from untransfected cells and those transfected pCGT7-Zfp647 or the empty pCGT7 vector. The cells were fractionated into nuclear and cytoplasmic fractions, and the nuclear extracts were separated into their soluble and insoluble components (Section 2.4.2). The extracts were then western blotted using an _-T7 antibody. No band was detected by the _-T7 antibody in the untransfected cells (Figure 3.16A), and no band corresponding to T7-Zfp647 was detected in the cells transfected with pCGT7 (Figure 3.16A). When prepared using 100mM NaCl, the vast majority of T7-Zfp647 is found in the insoluble nuclear extract, which is also the case at 200mM NaCl. This contrasts with the endogenous protein, which is soluble at 100mM NaCl (see Figure 3.11). I therefore tested the extracts made using 300-500mM NaCl, and found that only at 500mM NaCl is a significant fraction of the tagged protein soluble, and the majority is still found in the insoluble nuclear extract.

The zinc fingers of Zfp647 were also cloned into the pCGT7 vector (creating the pCGT7-Zfp647-ZF construct), to see whether their expression without the KRAB domain would yield a soluble protein, as it has been previously demonstrated that KRAB domains are highly insoluble (Tan *et al*, 2004a). Soluble and insoluble nuclear extracts, as well as whole cell extracts, were made from cells transfected with either the pCGT7 or pCGT7-Zfp647-ZF vector at a NaCl concentration of 200mM and western blotted with _-T7 antibody (Figure 3.16B). No signal is seen from the extracts of cells transfected with pCGT7, but a band of ~35kDa is seen in the whole cell extract from cells expressing T7-Zfp647-ZF. The zinc finger domain of Zfp647 is predicted to have a molecular weight of 41kDa. The T7-Zfp647-ZF protein is only seen in the insoluble fraction of the nuclear extract (lane 6 compared with lane 5), therefore removal of the KRAB domain from this protein does not increase its solubility.

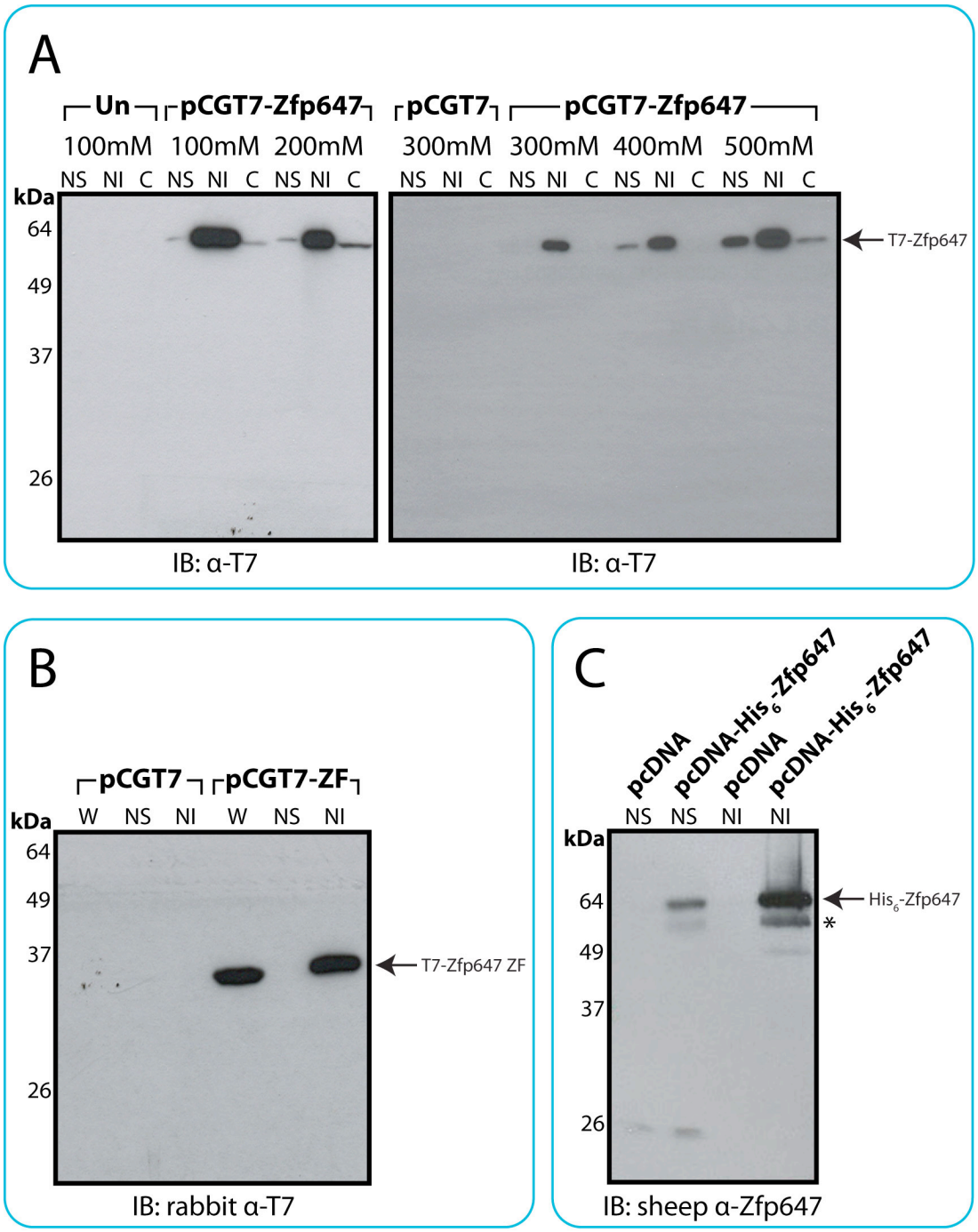


Figure 3.16. Expression of tagged Zfp647 constructs in mammalian cells. A) Western blot of soluble and insoluble nuclear, and cytoplasmic extracts of untransfected NIH/3T3 cells (Un), or cells transfected with pCGT7 or pCGT7-Zfp647, probed with mouse α-T7 antibody. Extracts were prepared using 100-500mM NaCl. B) Western blot of whole cell, nuclear soluble or nuclear insoluble extracts made from NIH/3T3 cells transfected with pCGT7 or pCGT7-Zfp647 ZF (pCGT7-ZF), probed with mouse α-T7 antibody. Extracts were made using 100mM NaCl. C) Western blot of soluble and insoluble nuclear extracts made from cells transfected with pcDNA3.1(+) (pcDNA) or pcDNA-His_{6}-Zfp647, probed with sheep α-Zfp647. Cells were made using 100mM NaCl. Possible degradation product is highlighted by an asterisk. NS: soluble nuclear extract; NI: insoluble nuclear extract; C: cytoplasmic extract; W: whole cell extract.

Cells were also transfected with either pcDNA or the pcDNA-His $_6$-Zfp647 construct. Nuclear extracts were made from these cells and again separated into their soluble and insoluble fractions. When western blotted with the rabbit _-Zfp647 antibody, the signal from the exogenously expressed protein is visible after a short exposure of the western blot (Figure 3.16C), whereas the endogenous protein appeared after a longer exposure (data not shown). Thus although the majority of His_6-Zfp647 is found in the insoluble fraction of the nuclear extract, the quantity of exogenous protein in the soluble fraction is greater than that of the endogenous protein.

It may have been possible to carry out a co-immunopreciptation experiment using the mammalian-expressed tagged Zfp647 proteins, following dialysis of T7-Zfp647 extracts from a high to low salt concentration, mixing these extracts together and immunoprecipitating with an antibody against one of the tags. However, I decided to attempt to express Zfp647 in a different manner, using an *in vitro* expression kit.

3.6.2. Immunoprecipitations of *In Vitro* Transcribed and Translated Zfp647

In vitro expression of mammalian proteins is a widely-used method, whereby proteins encoded in a plasmid can be transcribed and translated using normal amino acids to generate tagged proteins, as would occur in mammalian cell lines or bacterial cultures. There is also the possibility of including [^{35}S]-methionine into the reaction, labeling the proteins with a radioactive mark. Two differently tagged or labelled proteins can be incubated together after their production, and co-immunoprecipitated as with proteins expressed in tissue culture cells. This method has the additional benefit of ruling out any third party in the interaction, as using mammalian protein extracts may include factors that bridge between the two proteins being tested. On the negative side, proteins produced in this way will not be post-translationally modified, which would be critical if this is a requirement for the interaction to occur between the endogenous proteins. However, this method of expressing proteins has been successfully used in the study of KRAB zinc-finger

protein interactions previously (Tan *et al*, 2004a and b), therefore I decided to use it to try to detect the self-interaction of Zfp647.

3.6.2.1. Cloning, Expression and Verification of pGBKT7-Zfp647 and pGADT7-Zfp647

The full-length sequence of Zfp647, minus the start codon, was cloned into the yeast two-hybrid vectors pGBKT7 and pGADT7. As well as encoding the GAL4 activation and DNA-binding domains, pGBKT7 and pGADT7 encode c-myc and HA tags respectively, thus proteins cloned into them will be N-terminally tagged accordingly. The transcription of the tagged protein is under the control of a T7 promoter, therefore I expressed them using the Promega TNT T7 Quick for PCR DNA kit. Primers were designed against the pGBKT7-Zfp647 and pGADT7-Zfp647 plasmids that bound upstream of the c-myc or HA tags at the 5' end of Zfp647, and to the 3' end of the Zfp647 sequence itself, including the stop codon (Table 2.4). These primers were used to amplify the tagged Zfp647 sequence. As a control, the primers were also used in a PCR with the empty pGBKT7 and pGADT7 vectors. This should yield no PCR product, as the 3' primer binds to the Zfp647 sequence, thus amplification cannot occur when the vector is empty. The control PCR is identical to that used to generate Zfp647 in every other way, and can be used as a negative control in an immunoprecipitation where the only element missing is the tagged Zfp647 protein.

The pGBKT7, pGBKT7-Zfp647 and pGADT7-Zfp647 PCR products were used in TNT reactions, with [^{35}S]-methionine present to radio-label the proteins, and with unlabelled methionine to check the background signal from the plasmids (Section 2.7.1; Figure 3.17A). Both the unlabelled TNT reactions gave no signal, whereas the [^{35}S]-labelled pGBKT7 reaction produced a faint band at approximately 70kDa (Figure 3.17A, unfilled arrow), which is also present in the TNT products from the Zfp647 vectors. Both the pGBKT7-Zfp647 and pGADT7-Zfp647 TNT reactions produce a strong band at 80 kDa. This band is 20kDa larger than expected, which is

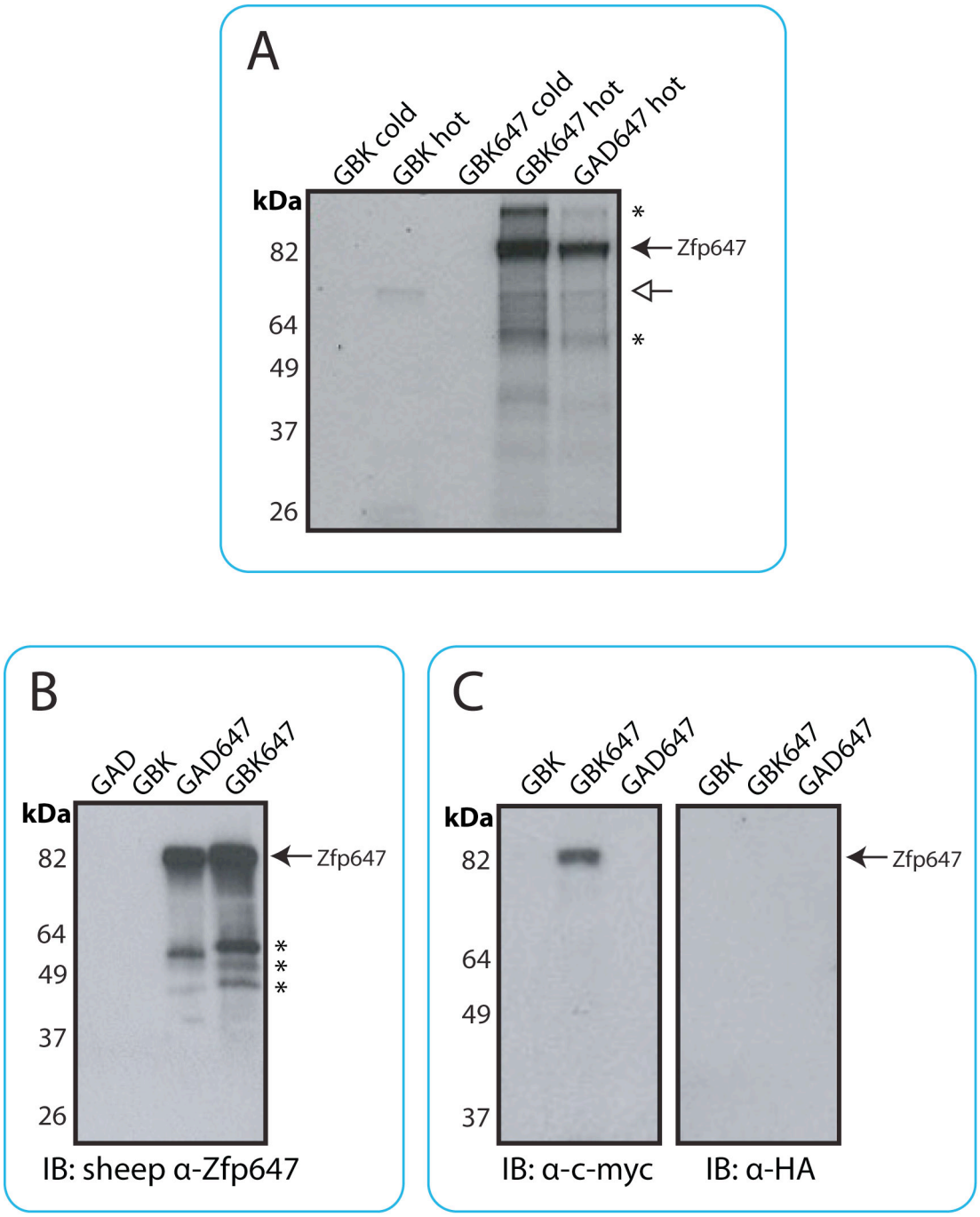


Figure 3.17. *In vitro* expression of Zfp647 by PCR TnT. GBK and GAD signify products from TnT reactions using PCRs of the empty pGBKT7 and pGADT7 vectors as template; GBK647 and GAD647 signify products from TnT reactions using the pGBKT7-Zfp647 and pGADT7-Zfp647 vectors as the template for the input PCRs. 5% of a 50µl TnT reaction was run out for each lane. A) Autoradiograph of [^{35}S]-labelled ('hot') and unlabelled ('cold') TnT products run out by PAGE. B) Western blot of cold TnT products probed with sheep α-Zfp647 antibody. C) Western blot of cold TnT products probed with rabbit α-c-myc and rabbit α-HA antibodies. Filled arrows indicate Zfp647; unfilled arrow indicates non-specific band from pGBKT7 plasmid; asterisks indicate unknown non-specific bands.

due to the fact that the TNT products were run on a pre-cast gel using a MOPS buffer system, which makes proteins appear to be 10-20kDa greater in mass than they are. The 80kDa band therefore probably represents the TNT-expressed Zfp647 protein. Two other relatively strong signals are seen in the [^{35}S]-labelled Zfp647 lanes at 110kDa and 60kDa (Figure 3.17A, highlighted with asterisks).

To verify the expression of the KRAB zinc-finger protein, the TNT reactions with both empty vectors and those carrying Zfp647 were repeated with unlabelled methionine, and the products were immunoblotted with sheep _-Zfp647 (Figure 3.17B). Both empty vectors gave no signal, whereas the antibody detected the same 80kDa band seen in Figure 3.17A in the TNT products of the vectors carrying Zfp647 (Figure 3.17B, filled arrow). The antibody also detected the 60kDa band seen previously (Figure 3.17B, unfilled arrow), suggesting that this band may be a Zfp647 degradation product.

3.6.2.2. Co-Immunoprecipitation of TNT-expressed Zfp647

To carry out the co-immunoprecipitation experiment, it was only necessary for one of the TNT-expressed proteins to have a recognisable tag, as the co-immunoprecipitated protein could be labelled with [^{35}S]-methionine to allow its detection. Both the tagged Zfp647 proteins were checked for how well their tags were recognised. Unlabelled proteins from TNT reactions using the pGBKT7, pGBKT7-Zfp647 (c-myc-tagged) and pGADT7-Zfp647 (HA-tagged) PCR products were western-blotted with antibodies against c-myc and HA, shown in Figure 3.17C. Neither of the antibodies gave a signal from the empty pGBKT7 vector. pGBKT7-Zfp647 is detected at the expected size only by the c-myc antibody, demonstrating that this protein was correctly tagged. However, pGADT7-Zfp647 did not give a band with either antibody, suggesting that either the HA-tag on this protein has not folded properly, or the _-HA antibody does not work.

[^{35}S]-labelled T NT products from both the empty vectors and those carrying Zfp647 were then subjected to immunoprecipitation by the antibody specific to their protein tag (Figure 3.18A). The translated proteins were incubated with sepharose beads bound by either _-c-myc or _-HA antibody, and either 5% of the input, or the entire immunoprecipitation was run out by PAGE and visualised by autoradiography. pGBKT7-Zfp647 is efficiently immunoprecipitated by the c-myc antibody, as the immunoprecipitation signal is stronger than that of the input, whereas pGADT7-Zfp647 is not pulled down to the same degree by the antibody against HA (Figure 3.18A). The proteins were also immunoprecipitated with the other antibody, i.e. pGBKT7-Zfp647 was immunoprecipitated with the HA antibody and *vice versa*: I found that neither protein immunoprecipitates with the antibody against the other tag (data not shown).

As c-myc-tagged Zfp647 was the better choice of protein to be immunoprecipitated, the pGBKT7-Zfp647 and pGBKT7 PCRs were expressed using unlabelled methionine. pGADT7-Zfp647 was expressed incorporating [^{35}S]-methionine. The co-immunoprecipitations were carried out using the protocol described by Gallo *et al*, which demonstrated the dimerisation of TNT-expressed *Drosophila* ADAR (Gallo *et al*, 2003). The pGADT7-Zfp647 TNT product was mixed with either the pGBKT7 or pGBKT7-Zfp647 reaction, and incubated for an hour at 37°C to allow complexes to form. The TNT reactions were immunoprecipitated with _-c-myc antibody bound to sepharose beads, and the immunoprecipitates run out on a polyacrylamide gel and analysed by autoradiography (Figure 3.18B). 5% of the reactions were run out for the input lanes; compared to these, and the signal from the immunoprecipitation of pGBKT7-Zfp647 by _-c-myc antibody (Figure 3.18A, lane 8), the signal from the co-immunoprecipitation lanes are quite weak. The amount of radiolabelled pGADT7-Zfp647 going into both experiments was equal, but the signal from the immunoprecipitation with pGBKT7-Zfp647 is much stronger than that from the immunoprecipitation with pGBKT7. It therefore appears that HA-tagged Zfp647 is specifically co-immunoprecipitated by c-myc-tagged Zfp647, demonstrating that Zfp647 can self-interact.

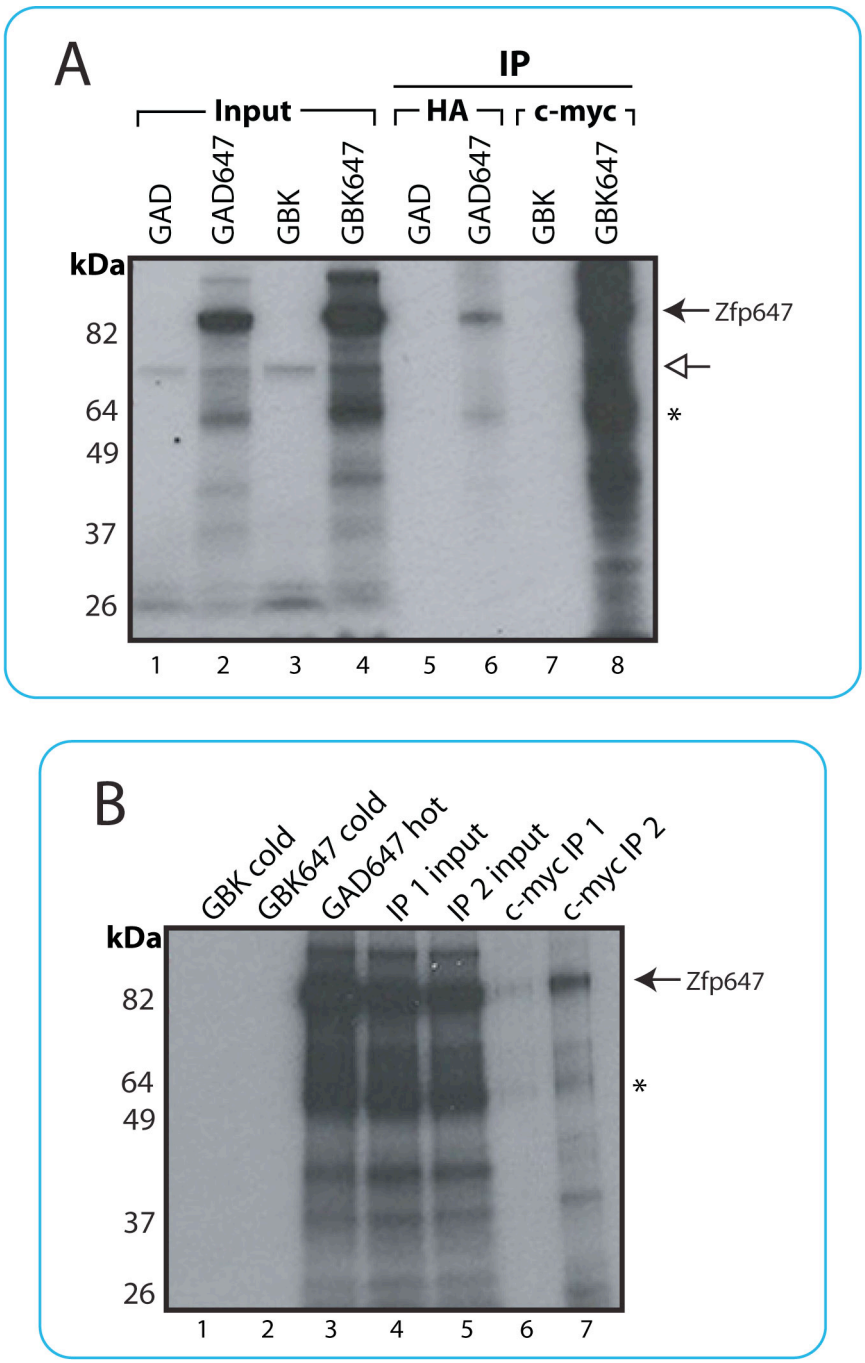


Figure 3.18. Immunoprecipitations of *in vitro*-expressed Zfp647. GBK and GAD signify products from TNT reactions using PCRs of pGBKT7 and pGADT7; GBK647 and GAD647 signify products from TNT reactions using pGBKT7-Zfp647 and pGADT7-Zfp647 PCRs. 5% of a 50µl TNT reaction was run out for each input; the entire immunoprecipitation was run out for IP lanes. A) Autoradiograph of the immunoprecipitation (IP) of [^{35}S]-labelled GAD, GAD647, GBK and GBK647 TNT products by either α-HA or α-c-myc antibody. B) Autoradiograph of co-immunoprecipitation of [^{35}S]-labelled GAD647 by unlabelled (cold) GBK or GBK647 TNT products. Lanes 1-3 show inputs before mixing; lanes 4 and 5 show the combined inputs into the immunoprecipitations; IP1: co-immunoprecipitation of GAD647 with GBK; IP2: co-immunoprecipitation of GAD647 with GBK647. Filled arrows indicate Zfp647; unfilled arrow indicates non-specific band from plasmid backbone; asterisks indicate non-specific bands.

3.7. Discussion

3.7.1. Zfp647 is Capable of Homo-Oligomerisation

Many KRAB zinc finger proteins have been shown to be capable of interacting with factors other than KAP-1, possibly giving them extra functions in the cell as well as regulating the transcription of their gene targets. In this chapter, I have demonstrated that on top of binding to KAP-1, Zfp647 interacted with itself in a yeast two-hybrid assay. To verify this self-interaction, I expressed tagged Zfp647 in both bacterial and mammalian cells, with the intention of using these proteins in a co-immunoprecipitation experiment. However, it appears that overexpressed Zfp647 is highly insoluble (Figures 3.15 and 3.16). It has previously been shown that the zinc fingers of a KRAB zinc finger protein are more soluble than the full-length protein (Tan *et al*, 2004a), however I found that the zinc fingers of Zfp647 were not soluble under normal conditions (Figure 3.16B). Therefore I carried out co-immunoprecipitation experiments with *in vitro*-expressed Zfp647 proteins. The signal from the co-immunoprecipitation of Zfp647 was fairly weak, however it did show that Zfp647 was specifically co-immunoprecipitated in the presence of another Zfp647 moiety (Figure 3.18B). I intended to carry out a positive control using KAP-1 to compare the signal from the Zfp647-Zfp647 co-immunoprecipitation to that of Zfp647 with a known interacting protein; however, I had difficulty cloning the full-length KAP-1 sequence into the yeast two-hybrid vectors, and thus did not manage to carry this experiment out.

I have shown that the self-interaction of Zfp647 occurs through the protein's zinc fingers, rather than the KAP-1-interacting KRAB domain. The ability to homo-oligomerise has only been demonstrated for one other KRAB zinc finger protein, ZBRK1, which also self-interacts through its C-terminal domain (Tan *et al*, 2004b). Homo-tetramerisation of ZBRK1 was shown to increase transcriptional repression; it is therefore possible that formation of Zfp647 complexes may also be important for its repressive activity, and that many of the other KRAB zinc finger proteins behave in a similar fashion. We have previously shown that the localisation of Zfp647 is

altered upon the removal of its zinc finger domain, as gene-trapped Zfp647 mislocalises to the heterochromatin (Briers *et al*, 2009). Thus it could be that the homo-oligomerisation ability of Zfp647, rather than its binding to DNA, is responsible for its localization to the KAKA foci. If we were able to express a Zfp647 construct in tissue culture cells that showed the same localization as the endogenous protein, we could mutate the protein-protein interaction domain to assess whether this affects the protein's recruitment to the KAKA foci.

3.7.2. Recruitment of Zfp647 to KAKA Foci is Cell-Type Dependent

Dr Heidi Sutherland has previously detailed its localisation of Zfp647 in mouse ES cells (Briers *et al*, 2009); in this thesis I have looked at Zfp647's localisation in NIH/3T3 and pMEF cells. I have found that unlike differentiated ES cells, NIH/3T3 cells do not have KAKA foci. Although the pattern of Zfp647 staining in these cells is slightly granular, they do not show the same distinct Zfp647- and KAP-1-containing foci seen in differentiated ES cells. In comparison, pMEFs do appear to have KAKA foci. These foci were seen in approximately 15% of cells. Each of the cell lines tested has been embryonically derived, and both the NIH/3T3 cells and pMEFs are fibroblasts; however, only the NIH/3T3 cells are immortalized, which may in some way be responsible for the differing behaviour of Zfp647 in this cell line. Alternatively, KAKA foci may be cell type-specific: certain factors promoting the formation of KAKA foci could be lacking in NIH/3T3 cells, or Zfp647 may be carrying out a different function in these cells. Therefore finding Zfp647 target genes and endogenous protein binding partners may shed light on why the protein gives different staining patterns in these cell lines. In the future, it may be of interest to look at the nuclear localisation of Zfp647 in tissue samples from regions of the mouse that it is expressed in, such as neural or heart tissue (Briers, 2005). We could then see whether the staining pattern seen in tissue samples matches that seen in ES and pMEF cells, or NIH/3T3 cells, and also whether KAKA foci are present *in vivo*. Unlike the localisation pattern of KRAZ1 and KRAZ2 (Matsuda *et al*. 2001), we have not seen Zfp647 localise to the pericentric heterochromatin in ES, pMEF or

NIH/3T3 cells. It could therefore be the case that different KRAB zinc finger proteins show different localisation patterns, according to their roles in the cell, or that over-expressed KRAB zinc finger fusion proteins mis-localise in the nucleus, as we have demonstrated with GFP-tagged versus wild-type Zfp647 (Figures 1.9 and 1.10; Briers *et al*, 2009).

3.7.3. Recruitment to KAKA Foci is KRAB Zinc Finger Protein-Specific

I identified 20 KRAB zinc finger proteins that interact with Zfp647 in my yeast two-hybrid screen, and thus could possibly co-localise with this protein at the KAKA foci in ES or pMEF cells. The localisation pattern of the majority of these proteins could not be analysed, as they have not had antibodies raised against them. This is not surprising, given the high levels of homology between proteins belonging to the same gene cluster (e.g. the Rslcan proteins), and the fact that so few KRAB zinc finger proteins have had a definite function assigned to them, thus most of these proteins have not been studied in great detail. I therefore raised an antibody against one of the KRAB zinc finger proteins identified in my yeast two-hybrid screen, Rbak. The antibody was raised in Guinea pig, to allow immunofluorescence studies with other antibodies raised in rabbit, mouse or sheep. Despite the pre-immune serum recognizing a nuclear band in cell extracts corresponding to Rbak (Figure 3.9), both the serum and the purified antibody failed to stain the nuclei of NIH/3T3 cells (Figure 3.10). However, the unpurified serum also detected a mainly cytoplasmic band by western blot (Figure 3.9), therefore it may be possible that this band is in fact Rbak, and the protein is cytoplasmic. To ascertain where Rbak localises to within the cell, it may be possible to inhibit its expression by a method such as RNAi, and see whether the nuclear or cytoplasmic western blot band disappears.

We have received antibodies against the KRAB zinc finger proteins NT2 (Tanaka *et al*, 2002), Zfp37 and Zfp90 (Payen *et al*, 1998), allowing us to analyse whether these proteins also co-localise with KAP-1 at KAKA foci. Dr Heidi Sutherland found that, like Zfp647, NT2 co-localised with KAP-1 in differentiated ES cells (Briers *et*

al, 2009). I have stained NIH/3T3 cells with the other two KRAB zinc finger protein antibodies, and found that these proteins do not form distinct nuclear foci; likewise, Dr Heidi Sutherland has obtained the same results from differentiated ES cells. This suggests either that not all KRAB zinc finger proteins localize to KAKA foci, or that their recruitment to these structures, if present in the cell, is cell type-specific. Given the number of KRAB zinc finger proteins in the mouse genome, it is not surprising to find that they do not all show the same pattern of staining, as one would predict them to be involved in a wide range of pathways.

3.7.4. Zfp647 Interacts with ARD1, Another Tripartite Motif Protein

One of Zfp647's interactions that was intriguing was that with ARD1. Although only one yeast colony out of the 800 in my screen encoded this protein, it is of interest as it is a member of the tripartite motif family, and has previously been shown to interact with the human Zfp647 homologue ZNF250, also in a yeast two-hybrid screen (Rual *et al*, 2005). ARD1 was the only tripartite motif protein reported to interact with ZNF250 in this screen; likewise, ZNF250 was the only KRAB zinc finger protein found to interact with ARD1 in the same screen. Although it has been pubished that ARD1 and KAP-1 do not interact (Reymond *et al*, 2001), Rual *et al* reported that these two proteins can co-purify (Rual *et al*, 2005). Thus it may be of interest to research the relationship of these two tripartite motif proteins, to see if they do interact *in vivo*, to analyse their localisation (does ARD1 localise to the KAKA foci or PML-NBs, both of which contain other tripartite motif proteins?) and to determine whether Zfp647 can interact with tripartite motif proteins other than KAP-1.

3.7.5. Zfp647 Interacts with Jumonji Domain-Containing Proteins

I have found that Zfp647 interacts with two jumonji domain-containing proteins, Jmjd1c and Jarid1b. Jmjd1c was identified in my yeast two-hybrid screen, whereas I

discovered the interaction with Jarid1b by co-immunoprecipitating Zfp647 interactors from a NIH/3T3 nuclear extract (Figures 3.13 and 3.14). The jumonji domain has the ability to demethylate histones (Cloos *et al*, 2008). Jmjd1c has not been shown to possess histone demethylase activity to date, but Jarid1b can remove methyl groups from H3K4 (Christensen *et al*, 2007; Seward *et al*, 2007; Yamane *et al*, 2007), and has been associated with breast, prostate and skin cancer (Lu *et al*, 1999; Xiang *et al*, 2007; Roesch *et al*, 2008). Both interactions require further confirmation to prove that they can interact with Zfp647, for example by *in vitro* co-immunoprecipitation experiments using epitope-tagged Jmjd1c or Jarid1b. Another KRAB zinc finger protein, Zkscan17, has been shown to interact with a different jumonji domain-containing protein, JARID2 (Mysliwiec *et al*, 2007), therefore there may be a link between these two classes of proteins.

In my *in vitro* co-immunoprecipitation experiments, there was a very high amount of background from proteins non-specifically interacting with the sepharose beads, which can be problematic if it increases the number of proteins identified in the mass spectrometry stage, possibly masking true interactors (Trinkle-Mulcahy *et al*, 2008). The mass spectrometry experiments could therefore be repeated using a protocol such as SILAC (stable isotope labelling with amino acids in cell culture) (Ong *et al*, 2002; Schulze and Mann, 2004). Cells are grown in the presence or absence of amino acids labelled with heavy carbon or arginine isotopes. Extracts from these cells are then subjected to an immunoprecipitation experiment, whereby the unlabelled cells are immunoprecipitated with a control antibody (e.g. _-GFP) and the heavy-labelled cells are immunoprecipitated with an antibody against the protein of interest, in this case _-Zfp647. The two immunoprecipitations are then pooled, run out by PAGE, and subjected to mass spectrometry. When the mass spectrographs are compared, a protein present in both the normal and heavy-labelled samples gives two peaks, owing to the shift in molecular weight caused by the heavy isotopes. Any non-specific background signal should give the same sized intensity peak in both the heavy- and normally-labelled samples, whereas a protein that is specifically co-immunoprecipitated by the _-Zfp647 antibody should give a higher peak in the heavy-labelled sample than in the normal sample. As there are many combinations of

heavy isotopes that can be used in SILAC experiments, it would be possible to compare proteins that are co-immunoprecipitated in NIH/3T3 cells with those from pMEF or differentiated ES cells. This could give an insight into other proteins interacting with Zfp647 at the KAKA foci, as these proteins may only be co-immunoprecipitated in the pMEF or ES cell extracts. Alternatively, the protein interactors of Zfp647 could be compared using the SILAC technique in undifferentiated and differentiated ES cells, to analyse whether these alter as the localisation of the protein changes.

One of the most interesting aspects of the yeast two-hybrid screen was the identification of a number of proteins involved in the ubiquitylation process. These included four proteins with RING domains, which are predicted to be able to catalyse the ubiquitylation (or possibly sumoylation) of a substrate protein. I therefore began to look at whether Zfp647 is in any way post-translationally modified.

Chapter 4: Post-Translational Modification of Zfp647

4.1. Introduction

A number of reasons led me to investigate whether Zfp647 is post-translationally modified. Firstly, five proteins that are involved with the ubiquitin machinery, or that have been shown to function as ubiquitin E3 ligases, were pulled out of my yeast two-hybrid screen (Section 3.2.6). Protein ubiquitylation is most commonly associated with degradation, as many proteins are poly-ubiquitylated before being destroyed by the proteasome. In this case, chains of the 76 amino acid ubiquitin moiety are covalently attached to lysine residues of a protein. These are recognised by the 26S proteasome, which cleaves the target protein into fragments that can be recycled by the cell (Hershko and Ciechanover, 1998). Poly-ubiquitylation can occur routinely as part of the cell cycle (e.g. cyclins (Cohen-Fix *et al*, 1996)) or to remove mis-folded or aggregated proteins from the cell (e.g. the endoplasmic reticulum-associated degradation pathway (reviewed in Vembar and Brodsky, 2008)). Mono-ubiquitylation can also alter a protein's function, rather than signal its impending destruction (Welchman *et al*, 2005).

Secondly, two KRAB zinc-finger proteins were identified in a proteomics screen looking for proteins that are modified by the 101 amino acid, 11kDa small ubiquitin-like modifier (SUMO) in HeLa cells (Vertegaal *et al*, 2006). These proteins were ZNF317 (comprised of a KRAB domain and 13 zinc-finger motifs) and RBAK (Section 3.4). Like ubiquitin, SUMO is attached to lysine residues of target proteins. Although sumoylation has been linked to the protein degradation process (PML protein is poly-sumoylated prior to degradation when cells are treated with As2O3, discussed in Section 1.2.5.), this modification functions more commonly to alter the behaviour of a protein or its localization in the cell, typically by changing its interactions with other proteins. The finding that KRAB zinc finger proteins can be modified by SUMO may mean that within the cell there are different subsets of each protein, capable of carrying out different tasks. Sumoylation of the KRAB zinc finger proteins may also be required for their repressive function.

Finally, since starting this project it has been found that KAP-1 is also modified by SUMO at six sites within the C-terminal end of the protein (Mascle *et al*, 2007; Lee *et al*, 2007b). Sumoylation of KAP-1 by Ubc9 is directed by its PHD and bromodomains, and is necessary for the interactions between KAP-1, SETDB1 and Mi-2α (Figure 4.1; Ivanov *et al*, 2007). The binding of KAP-1 to a KRAB domain actually increases the sumoylation levels of KAP-1 (Mascle *et al*, 2007). Therefore sumoylation is already known to play a role in the KRAB-mediated transcriptional repression pathway, and KAP-1 itself has been shown to possess E3 ligase activity.

4.1.1. The Ubiquitin-Like Proteins

There are three major ubiquitin-like modifying proteins found in mammalian cells: ubiquitin, SUMO and Nedd8. These proteins do not show a high degree of similarity at the amino acid level (ubiquitin and SUMO-1 are 18% identical; ubiquitin and Nedd8 are 58% identical; SUMO-1 and Nedd8 share 19% identity), however they each possess a conserved structure, the ubiquitin fold (Welchman *et al*, 2005). The method of conjugation for each of the ubiquitin-like proteins is also highly similar (Figure 4.2). The proteins are first activated to allow them to be covalently attached to the substrate protein; in the case of SUMO and Nedd8, C-terminal amino acid residues are removed from the protein, revealing the reactive diglycine motif. All three of the modifiers are activated by E1 activating enzymes, which results in their attachment to the E1 enzyme (Hershko and Ciechanover, 1998). The ubiquitin-like protein is then passed to an E2 ubiquitin-conjugating enzyme (Ubc). Mammals possess at least 30 different Ubcs specific for ubiquitin, whereas SUMO and Nedd8 are only transferred to one or two enzymes specific for their modification: Ubc9 for SUMO (Johnson and Blobel, 1997); UbcH12 and Ube2f for Nedd8 (Huang *et al*, 2009). All three of the modifying proteins can then be transferred to their intended substrate directly from the E2 enzyme. For ubiquitin and Nedd8, this process is catalysed by an E3 ligase. The E3 ligase can interact with both the substrate and the E2 enzyme, and thus acts as a substrate recognition factor for the E2 enzyme (Lorick *et al*, 1999). RING domain containing proteins function in

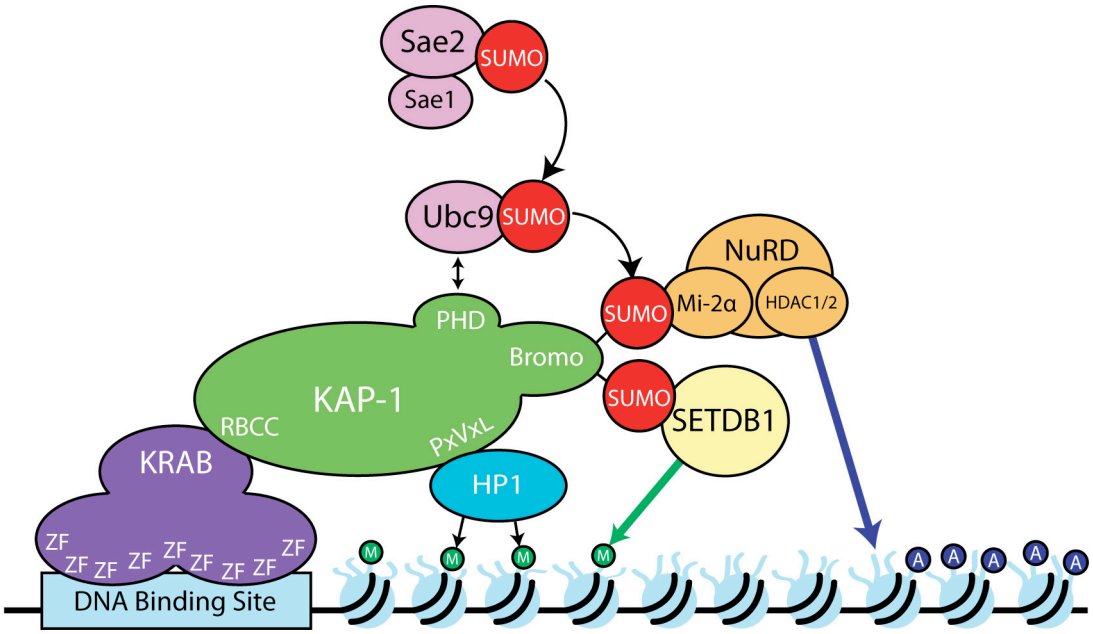


Figure 4.1. KAP-1 sumoylation in the KRAB repression pathway. The PHD of KAP-1 directs Ubc9 to sumoylate its bromodomain. The interaction of KAP-1 and a KRAB domain increases the level of KAP-1 sumoylation. Mi-2α and SETDB1 can only bind to the sumoylated form of KAP-1. NuRD and SETDB1 can then catalyse the deacetylation of histones and the methylation of H3K9 in the proximity of the KRAB binding site. Figure adapted from Ivanov *et al*, 2007.

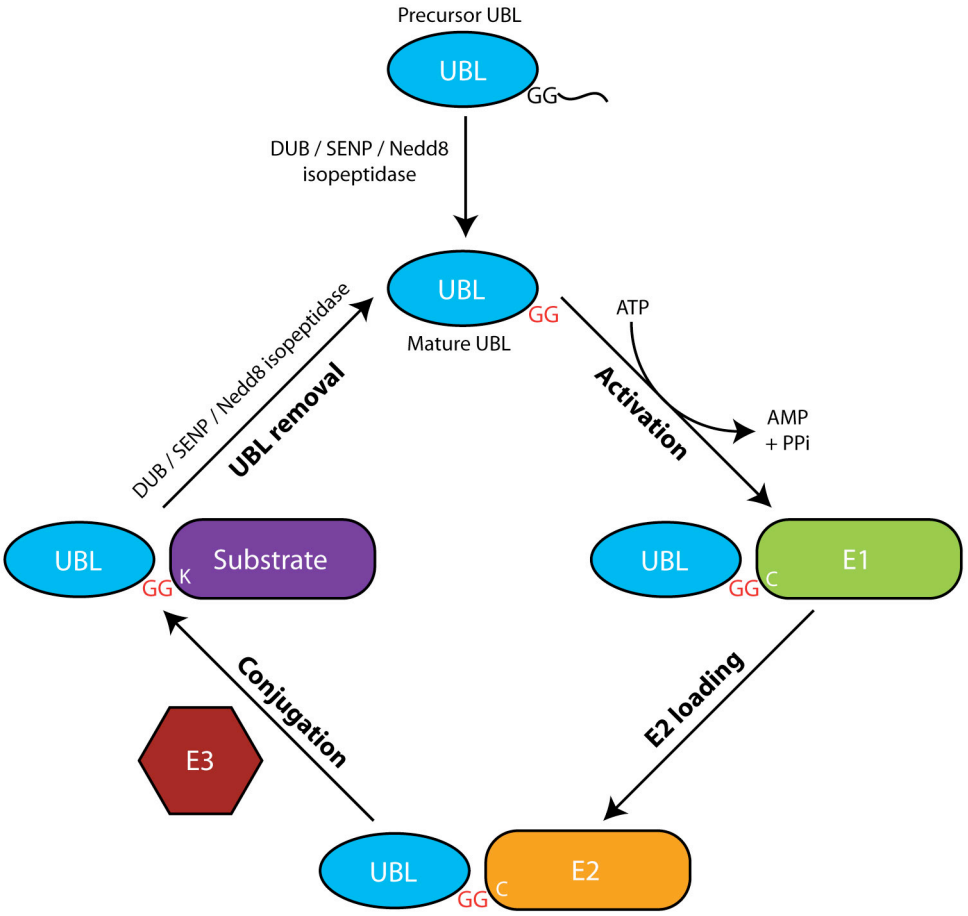


Figure 4.2. Conjugation cycle of the ubiquitin-like modifiers. Precursor ubiquitin-like modifiers (UBLs) are processed to generate the mature form of the protein, which possesses a reactive diglycine (GG) motif, by deubiquitinating enzymes (DUBs), sentrin-specific proteases (SENPs) or Nedd8 isopeptidases. The UBL is then activated and attached to the E1 enzyme, passed to an E2 conjugating enzyme, and transferred to the substrate protein, which often requires the action of an E3 ligase. The DUBs, SENPs and Nedd8 isopeptidases also catalyse the removal of the UBL from its substrate, generating a mature, free, UBL.

this manner, whereas HECT (homologous to the E6-AP carboxyl terminus) domain-containing E3 ligases actually act as recipients of the activated ubiquitin molecule, which can then be transferred to the substrate (Scheffner *et al*, 1993). In the case of SUMO, Ubc9 can bind to a conserved motif found in the substrate protein, with the consensus of ψKxD/E, where ψ stands for a large, hydrophobic amino acid (isoleucine (I), leucine (L), methionine (M), phenylalanine (F) or valine (V)), and x represents any amino acid (Rodriguez *et al*, 2001). Therefore the sumoylation of many proteins does not require the presence of an E3 ligase. However, not all proteins that encode the consensus sequence are sumoylated, and not all proteins that are sumoylated encode the consensus sequence (Vertegaal *et al*, 2006). Thus proteins belonging to this latter class are sumoylated with the aid of an E3 ligase to direct Ubc9 to the substrate; KAP-1's PHD and bromodomains function as its E3 ligase (Ivanov *et al*, 2007).

Modification by each of the ubiquitin-like proteins is reversible. Deubiquitinating enzymes (DUBs), sentrin-specific proteases (SENPs) and Nedd8 isopeptidases catalyse the removal of ubiquitin, SUMO and Nedd8. These enzymes function continuously, therefore the pool of modified forms of an endogenous protein that is susceptible to dynamic modification turnover can be undetectable in cell extracts (Li *et al*, 2007; Martin *et al*, 2007; Rajan *et al*, 2005)

4.1.2. Substrates of the Ubiquitin-Like Modifiers

The major substrates of Nedd8 are the cullin family, which themselves function as scaffolding proteins for ubiquitin E3 ligases, as they can interact with RING domain-containing proteins to aid in the recruitment of the ubiquitin activating E2 enzymes (Welchman *et al*, 2005). The modification of the cullins by Nedd8 can prevent their binding to inhibitory proteins, thus neddylation can influence the activity of ubiquitin E3 ligases. By contrast, many different classes of protein are modified by ubiquitin and SUMO. In some cases, these two modifications compete for the same lysine in the substrate's sequence. For example, ubiquitylation of IκB at

lysine 21 results in the protein's degradation, whereas sumoylation of the same residue stabilizes the protein (Welchman *et al*, 2005). Ubiquitylation is often associated with the endocytosis pathway, and a number of transcriptional regulators are also modified by ubiquitin. The bulk of SUMO substrates are also involved in the transcription process (Girdwood *et al*, 2004). Factors such as Sp3, c-Jun, the histone deacetylases HDAC1 and HDAC4, histone H4 and p300 can all be sumoylated, and their modification in this manner is associated with an increase in transcriptional repression (Girdwood *et al*, 2004). As discussed in Section 1.3.5, KAP-1 is also subjected to modification by SUMO at six lysine residues (Lee *et al*, 2007b; Mascle *et al*, 2007). Sumoylation of KAP-1 is essential for its interaction with Mi-2_ and SETDB1 (Ivanov *et al*, 2007), and thus this modification enhances KAP-1's transcriptional repressor function.

4.1.3. Effect of Sumoylation on Nuclear Localisation

While the exact details of how SUMO alters the behaviour of some of its substrates is unknown, it is clear that there is an association between sumoylation and nuclear localization. In particular, there is a close link between sumoylated proteins and the PML-NBs. Homozygous Ubc9-knockout mice have disrupted PML-NBs (Nacerddine *et al*, 2005), and knockout of SUMO-1 also leads to a decrease in both PML levels and the number of PML-NBs seen in mouse cells (Evdokimov *et al*, 2008). Sumoylated PML is only observed at the PML-NBs, and not in the nucleoplasm (Ishov *et al*, 1999), and it has been argued that PML sumoylation is essential for the formation of PML-NBs, as $PML^{-/-}$ cells transfected with a mutant form of PML that cannot be sumoylated do not display these nuclear bodies (Zhong *et al*, 2000). The SIM domains of PML and other proteins that localize to the PML-NBs are also essential for maintaining PML-NB structure, as these allow sumoylated forms of PML and proteins such as Daxx to interact (Ishov *et al*, 1999; Lin *et al*, 2006; Shen *et al*, 2006). Additionally, the overexpression of SENP2 leads to the redistribution of several PML-NB proteins (Dellaire and Bazett-Jones, 2004). Therefore it has been proposed that PML-NBs can act as repositories for sumoylated

transcription factors and other proteins (Girdwood *et al*, 2004), and their composition could be continuously altered by the sumoylation and desumoylation of PML-NB-associated proteins.

Sumoylation is also important for the regulation of the nuclear pore complexes, which are responsible for the trafficking of certain proteins in and out of the nucleus. One of the major recipient of the SUMO modification in mammalian cells is RanGAP1 (Ran GTPase-activating protein 1). RanGAP1 activates Ran, itself a GTPase involved in the transport of RNA and proteins between the nucleus and cytoplasm (Bischoff *et al*, 1994). Sumoylation of RanGAP1 is required for its localization to the nuclear periphery, through its interaction with RanBP2, another protein situated at the periphery (Matunis *et al*, 1996; Mahajan *et al*, 1997). RanGAP1, RanBP2 and Ubc9 have been shown to be capable of forming a ternary complex (Reverter and Lima, 2005). In accordance with this finding, a small proportion of cellular Ubc9, which is generally found in the nucleoplasm, can be visualized at the nuclear pore complexes (Lee *et al*, 1998; Saitoh *et al*, 2002). The SUMO proteases SENP1 and SENP2 can be found at the nuclear envelope, as well as in PML-NBs (Best *et al*, 2002; Hang and Dasso, 2002; Zhang *et al*, 2002). Yeast that have lost either the SUMO activating enzyme Uba2 or the SUMO protease Ulp1 are unable to import proteins bearing a classic nuclear localization signal into the nucleus (Stade *et al*, 2002). Therefore sumoylation and desumoylation of nuclear pore proteins appear to play an important role in controlling the transport of certain factors in and out of the nucleus.

4.2. Modification of Zfp647

4.2.1. Detection of High Molecular Weight Forms of Zfp647

The ubiquitin-like modifications can be rapidly turned over *in vivo*, a process that can also continue following cell lysis. Therefore the modified forms of a protein often go undetected when analysed by western blot. To prevent the removal of modifications, cell extracts can be made in the presence of N-ethylmaleimide (NEM). NEM reacts with the thiol groups found in cysteine residues, causing their covalent modification. As the enzymes that remove ubiquitin, SUMO and Nedd8 tags from proteins have a reactive cysteine residue at their core, NEM prevents these enzymes from catalyzing the modification's removal.

To see if modified forms of Zfp647 could be detected, I analysed protein extracts prepared in the presence of NEM. NIH/3T3 cells were treated with 20mM NEM for one hour at 37°C, and whole cell extracts were prepared with buffers containing 20mM NEM. Cells were also treated with the proteasome inhibitor MG132 at a concentration of 10μM, and the same volume (15μl) of the MG132 solvent, DMSO. If Zfp647 is poly-ubiquitylated ahead of its degradation, the inhibition of the proteasome would lead to the accumulation of this modified form of the protein, thus distinguishing it from any other form of the protein seen. Extracts were run out by PAGE and western blotted with sheep _-Zfp647 antibody (Figure 4.3). The extracts made from DMSO- and MG132-treated NIH/3T3 cells show the major 65kDa band normally seen with this antibody, and non-specific bands can also be seen at ~45 and 50kDa. Treatment of the cells with NEM leads to the appearance of at least two other strong bands, at ~70 and 85kDa. The 85kDa band appears to be present at very low levels in extracts made in the absence of NEM, but incubation of the cells with this chemical increases the signal from this protein band so that it is present at approximately the same level as the main Zfp647 band. There are also bands visible between 70kDa and 85kDa. It therefore appears that Zfp647 is post-translationally modified in tissue culture cells. The absence of any extra bands in extracts made from cells treated with MG132 suggests that these modifications have not arisen due

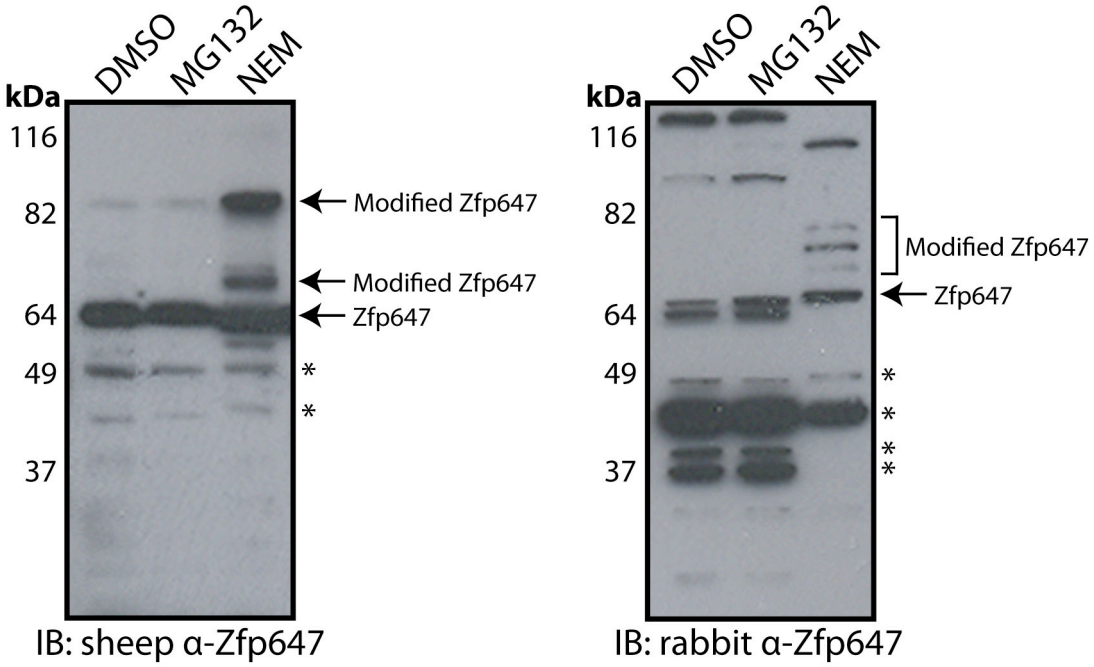


Figure 4.3. Detection of high molecular weight species of Zfp647. Western blots of whole cell extracts of NIH/3T3 cells treated with DMSO, 10μM MG132 or 20mM NEM. Extracts were run out by PAGE and western blotted with either sheep α-Zfp647 or rabbit α-Zfp647. Zfp647 and possible modified forms of Zfp647 are indicated with arrows; asterisks indicate non-specific bands.

to blocking of the degradation-mediated poly-ubiquitylation of Zfp647.

4.2.2. Verification of High Molecular Weight Zfp647 Species

To confirm that the high molecular weight bands are Zfp647, and not another protein species that is only present after NEM treatment, I decided to cross-test the two _-Zfp647 antibodies. First, to verify that the rabbit _-Zfp647 antibody could identify modified Zfp647, I probed the NIH/3T3 cell extracts made following treatment with DMSO, MG132 and NEM with this antibody (Figure 4.3). Rabbit _-Zfp647 detects the 65kDa Zfp647 band in all three extracts, and the 40kDa cytoplasmic band detected by this antibody (Figure 3.12) is also seen in each of the extracts. Whilst the DMSO- and MG132-treated cell extracts are identical in their pattern of staining, the pattern seen following NEM treatment is different. Three higher molecular weight bands are seen between 65kDa and 80kDa, one of which appears to be the same weight as that observed with the sheep antibody, and two high molecular weight species highlighted by the sheep _-Zfp647 antibody are replaced by one at 115kDa. Thus, except for the band at 70kDa, the proteins detected by the sheep _-Zfp647 antibody are different to those observed with the rabbit antibody; however, the staining by both antibodies is changed when NEM is present.

Despite the apparent differences in their specificity, I decided to test whether both antibodies can recognise the 70kDa Zfp647 band. To do this, I made nuclear extracts from NIH/3T3 cells prepared in the presence or absence of NEM, and immunoprecipitated Zfp647 using the rabbit antibody. I then ran the immunoprecipitates out by PAGE, and western blotted them using the sheep _-Zfp647 antibody (Figure 4.4). As seen previously, there is cross-reaction between the rabbit and sheep antibodies, as the IgG bands are visible at 60kDa. In the absence of NEM, sheep _-Zfp647 only detects the 65kDa Zfp647 band in the input lane, whereas when prepared with NEM, two further bands at ~75kDa and ~85kDa are seen in the input, similar to those in Figure 4.3. In the immunoprecipitation lane, in the absence of NEM, Zfp647 can be seen just above the IgG band, at ~65kDa. When

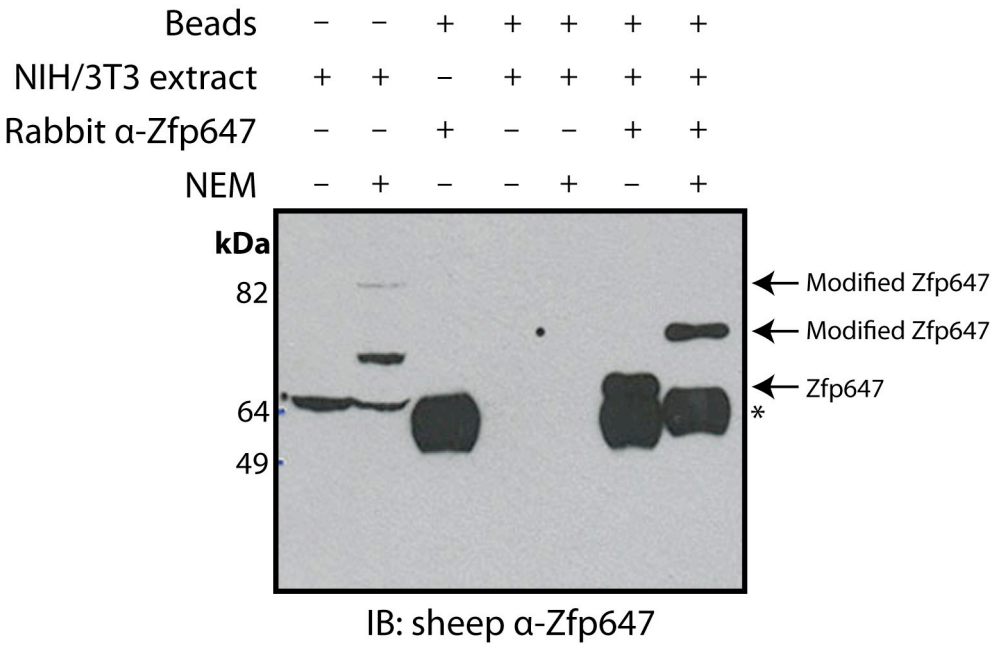


Figure 4.4. Immunoprecipitation of modified Zfp647. Western blot of Zfp647 immunoprecipitation from NIH/3T3 nuclear extracts, carried out with rabbit α-Zfp647, probed with sheep α-Zfp647. Immunoprecipitations were carried out in the absence or presence of NEM. Asterisks indicate IgG bands.

NEM is present, however, the strongest Zfp647 band seen is that at ~75kDa, corresponding to the modified protein. Thus both antibodies can recognize this protein, making it very likely that it is in fact a modified form of Zfp647.

4.2.3. Effect of KAP-1 on Zfp647 Modification

KAP-1 is known to direct its own modification by acting as an E3 ligase, binding to the E2-conjugating enzyme Ubc9 and directing it to the sites at which it can be sumoylated (Ivanov *et al*, 2007). It may be possible that KAP-1 could also direct Ubc9 to sumoylate a KRAB zinc finger protein that it is bound to. Alternatively, the RING domain of KAP-1 could function as a ubiquitin E3 ligase to modify a KRAB zinc finger protein.

To assess whether KAP-1 has any effect on Zfp647 modification, I chose to look at whether the higher molecular weight forms of Zfp647 were present in cells deficient in KAP-1. I received two PCC4 EC cell lines from Dr Daniel Wolf (Howard Hughes Medical Institute, New York), one of which is wild-type. The other cell line, named 111, encodes a small interfering RNA (RNAi) expression construct targeted against KAP-1, which has previously been shown to dramatically reduce KAP-1 protein levels when compared to PCC4 cells carrying a scrambled RNAi construct (Wolf *et al*, 2008). The PCC4 cells can be easily differentiated by the addition of RA, and I chose to look at undifferentiated cells and those treated with RA for 6 days in case the differentiation process affects either the level or activity of KAP-1. Whole cell extracts were made from both lines, and were western blotted with rabbit _-KAP-1 to verify the low level of KAP-1 in the 111 cells (Figure 4.5). KAP-1 is present at high levels in both the undifferentiated and differentiated extracts from wild-type cells. In comparison, KAP-1 is barely seen in the PCC4 111 cells. Down-regulation of the expression of KAP-1 over the course of differentiation has been previously described (Cammas *et al*, 2004; Wolf and Goff, 2007; Briers *et al*, 2009), but is only apparent in the –NEM lanes in Figure 4.5, possibly due to the long exposure of the western blot. The extracts were also blotted with an _-tubulin

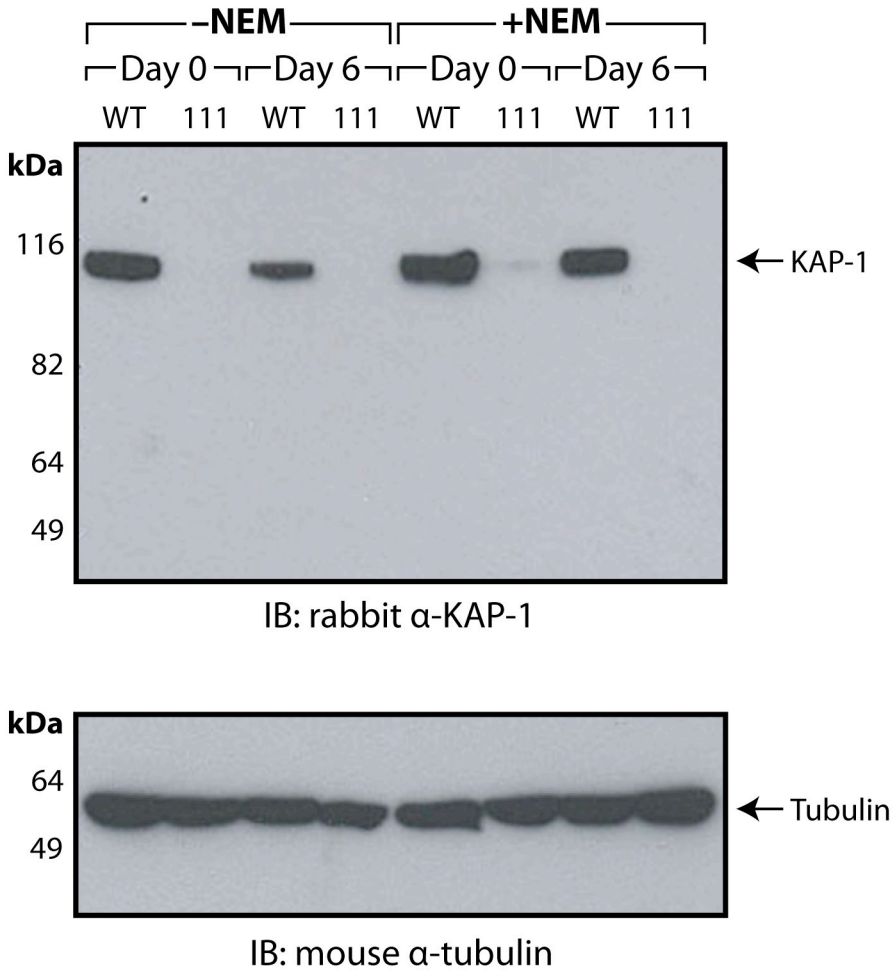


Figure 4.5. KAP-1 levels in PCC4 cells. Western blots of whole cell extracts made from PCC4 wild-type (WT) and 111 (encoding RNAi construct targeting KAP-1) cell lines in the presence and absence of NEM. Extracts were made from undifferentiated cells (Day 0) and those differentiated by RA treatment for 6 days (Day 6), and were western blotted with rabbit α-KAP-1 (upper panel) and mouse α-tubulin (lower panel).

antibody as a loading control (Figure 4.5). Despite the reports of KAP-1 modification by SUMO, extra bands corresponding to sumoylated KAP-1 were not seen in the extracts treated with NEM in the PCC4 cells, and I also did not observe modified KAP-1 in NIH/3T3 cells (data not shown). Thus, either the _-KAP-1 antibody used does not recognize modified KAP-1, KAP-1 is not modified in either of these cell lines, or the level of modified endogenous KAP-1 is too low to be visualised.

Western blots of the PCC4 cell extracts incubated with sheep _-Zfp647 antibody are shown in Figure 4.6. In the PCC4 111 cells, the signal corresponding to Zfp647 seems to be lower than that in the wild-type cells. This is seen in the extracts made both with and without NEM. Modified Zfp647 can be seen in each of the extracts made with NEM, at ~5kDa above the height of the main Zfp647 band, but the intensity of this band appears reduced, presumably due to the decreased overall level of Zfp647. The higher molecular weight form of Zfp647 can also be seen faintly in the undifferentiated PCC4 cell extracts from both the wild-type and 111 lines after a longer exposure (middle panel). These results suggest that the loss of KAP-1 does not have an effect on Zfp647 modification, but may in fact impact upon the overall levels of Zfp647 protein in the cell.

4.3. Is Zfp647 Modified by a Ubiquitin-Like Protein?

4.3.1. Analysis of Zfp647 Sequence

There does not appear to be any consensus sequence for the sites of attachment of ubiquitin and Nedd8. However, SUMO tends to be attached to lysines that are part of the consensus sequence ψKxD/E (Rodriguez *et al*, 2001), therefore I analysed the amino acid sequence of Zfp647 for the presence of this sequence. Zfp647 has 46 lysine residues, but none of these are part of a ψKxD/E consensus (Figure 4.7). Two of the lysines, at positions 99 and 135, are preceded by a small hydrophobic residue,

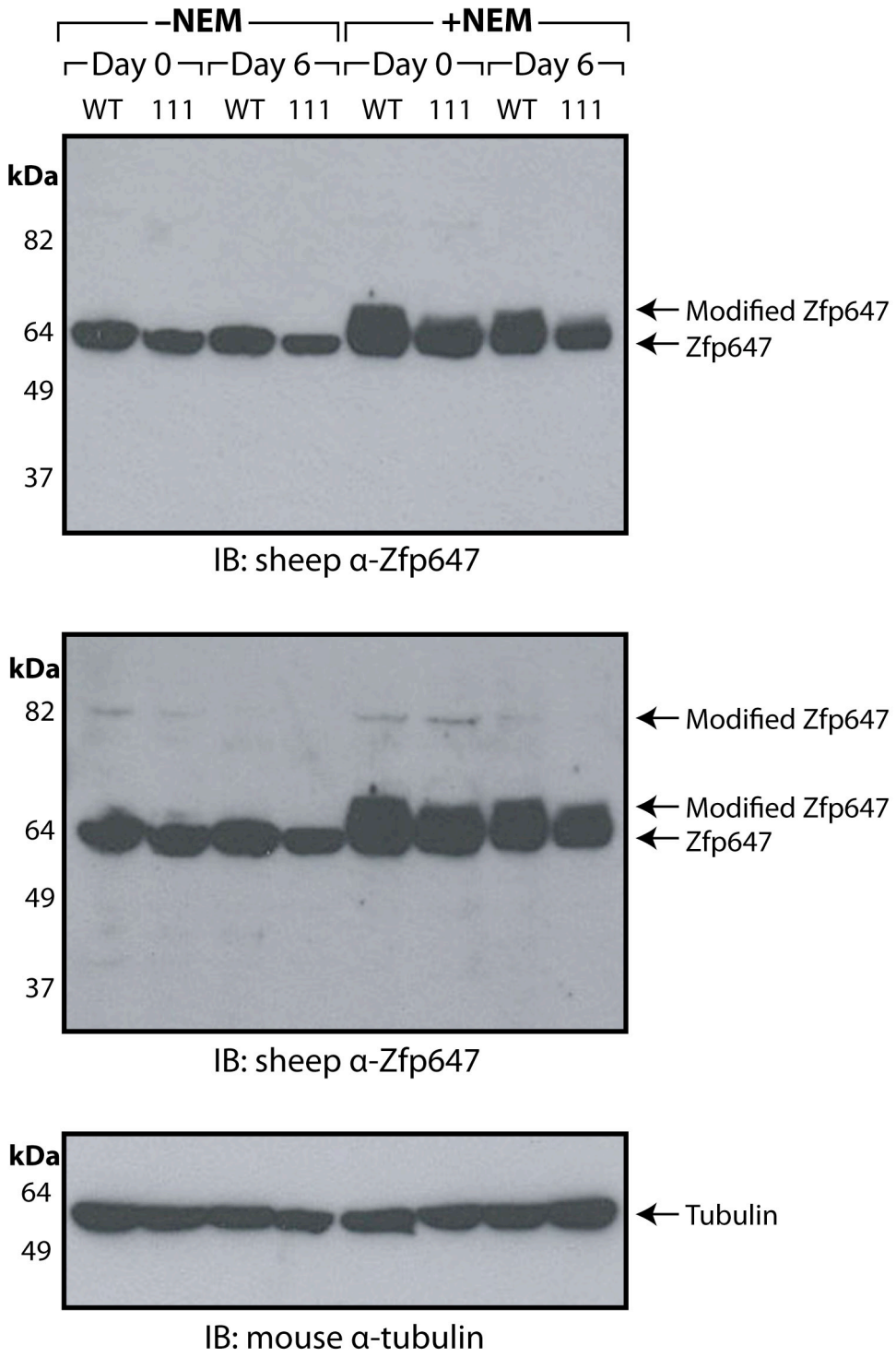


Figure 4.6. Zfp647 modification in PCC4 cells. Western blot of whole cell extracts made from PCC4 wild-type (WT) and 111 (encoding RNAi construct targeting KAP-1) cell lines in the presence and absence of NEM. Extracts were made from undifferentiated cells (day 0) and those differentiated by RA treatment for 6 days (day 6), and were western blotted with sheep α-Zfp647 (upper and middle panels) and mouse α-tubulin (lower panel). Middle panel is a longer exposure of the same blot as the upper panel.

```
  1 maaagllplp aapqakvtfe dvavllsqee warlgpaqrg lyrhvmmety
 51 gnvvslglpg skpvvisqle rgedpwvldg qgtelsqslg sdhseckake
101 enqntdlnvp plisdeasat ltetplrkva eeryktepkv cpspkpigpq
151 nahglnpsvp varpqtapsv arpyiciecg kcfgrsshll qhqrihtgek
201 pyvchvcgka fsqssvlskh rrihtgekpy ecnecgkafr vssdlaqhhk
251 ihtgekphec lecgkaftql shliqhqrih tgerpyvcpl cgkafnhstv
301 lrshqrvhtg ekphgcsecg ktfsvertll qhqrvhtgek pytcsecgka
351 fsdrsvliqh hnvhtgekpy ecsecgktfs hrstlmnher ihtqekpyac
401 yecgkafvqh shliqhqrvh tgekpyvcge cghafsarrs liqherihtg
451 ekpfqctecg kafslkatli vhlrthtgek pyecnscgka fsqysvliqh
501 qrihtgekpy ecgecgrafn qhghliqhqk vhkkl
```

Figure 4.7. Amino acid sequence of Zfp647. KRAB domain is highlighted in red; zinc finger motifs are highlighted in blue; lysine residues are highlighted in orange; and peptide sequences almost matching the sumoylation consensus sequence (small hydrophobic residue-K-any amino acid-D/E) are highlighed in green.

and followed by a glutamic acid residue at positions 101 and 137 (AKEE for the lysine at position 99; YKTE for the lysine at position 135; see Figure 4.7). Thus although Zfp647 does not have a consensus sumoylation motif, it does have two lysine residues that almost conform to the consensus. Although the four amino acid consensus sequence was initially thought to be essential (Sampson *et al*, 2001), proteins lacking this motif have been shown to be sumoylated *in vitro* (Vertegaal *et al*, 2006). Indeed, of the six sumoylation sites in KAP-1, only two, at positions 554 and 676, conform to the consensus motif (Lee *et al*, 2007b; Mascle *et al*, 2007).

4.3.2. Detection of Modification by Immunoprecipitation

Because the amino acid sequence of Zfp647 did not suggest whether this protein is subjected to sumoylation, I tried to ascertain what species was modifying this protein using biochemical means. Both ubiquitin and SUMO seemed likely candidates, due to the yeast two-hybrid results and the links between transcriptional repressors and SUMO. Furthermore, NEM is known to preserve the modification of substrates by the ubiquitin-like proteins. The molecular weights of the modified bands (5-10kDa and 20-25kDa greater than the main Zfp647 band) could possibly be due to the addition of one and two ubiquitin-like modifications, as these proteins weigh approximately 9kDa (ubiquitin and Nedd8) and 11kDa (SUMO). It has also been found that proteins modified by SUMO on different lysine residues can actually run at different mobilities when analysed by western blot. Mascle *et al* showed that KAP-1 mono-sumoylated at either lysine 779 or 804 ran with a distinctly higher mobility than KAP-1 sumoylated at lysine 554 (Mascle *et al*, 2007). Therefore these two different modified forms of Zfp647 could be due to single modifications at different sites within the protein's structure.

There are three forms of SUMO, SUMO-1, -2 and -3. SUMO-2 and SUMO-3 show a very high level of homology to each other, differing by only three amino acids, and are considered to function in exactly the same manner (Hay, 2005). Therefore these proteins will be jointly referred to as SUMO-2/-3. SUMO-1 shares

50% identity with SUMO-2/-3, and is known to behave independently of the two other SUMO family members. Screens for proteins modified by sumoylation have shown that some proteins are specific for which SUMO protein they are modified by, whereas some proteins can be modified by any of the SUMO proteins (Vertegaal *et al*, 2006).

I obtained antibodies against SUMO-1, SUMO-2, ubiquitin and Nedd8, with the intention of immunoprecipitating modified Zfp647 with an _-Zfp647 antibody, and western blotting the immunoprecipitated material with the antibodies against the ubiquitin-like modifiers. As three of these antibodies were raised in rabbit, I immunoprecipitated Zfp647 with sheep _-Zfp647 to avoid high levels of background in the western blot. The immunoprecipitations were carried out using NIH/3T3 whole cell extracts either with or without NEM. To test for the presence of modified Zfp647 in the immunoprecipitate, I first western blotted the pull-down material with rabbit _-Zfp647 (Figure 4.8). The input lanes show the effect of NEM on Zfp647: the main band detected increases in size by approximately 10kDa. The 65kDa Zfp647 band is visible in the immunoprecipitation above the IgG band. Modified Zfp647 is present in the +NEM immunoprecipitation, demonstrating that like rabbit _-Zfp647, the sheep _-Zfp647 antibody is capable of immunoprecipitating both unmodified and modified Zfp647. The immunoprecipitation carried out in the presence of NEM was then western blotted with each of the antibodies specific for the ubiquitin-like modifiers (Figure 4.9). The 75kDa modified Zfp647 band was not detected in the immunoprecipitated lanes. Therefore there was no indication from this experiment that Zfp647 was modified by SUMO, ubiquitin or Nedd8.

4.3.3. Transfection of Cells with Ubiquitin-Like Modifiers

To try to ascertain whether any of the ubiquitin-like proteins could be covalently attached to Zfp647, I decided to transfect cells with tagged forms of ubiquitin, SUMO and Nedd8. Endogenous proteins can be modified by exogenous forms of the

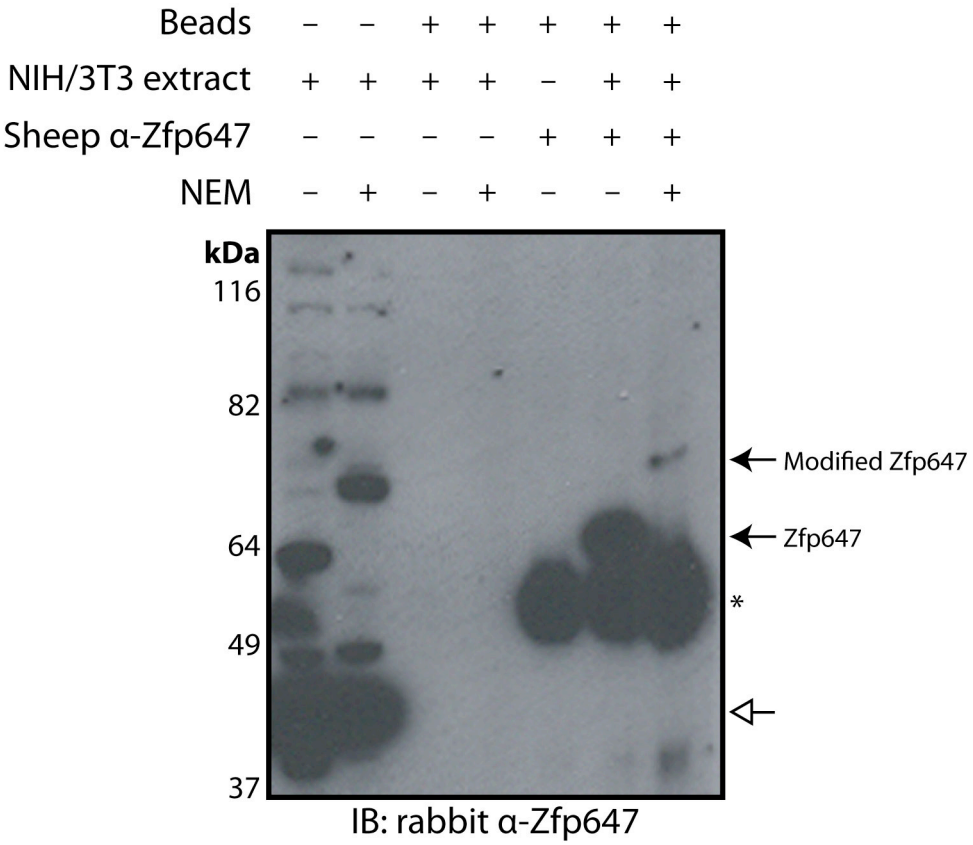
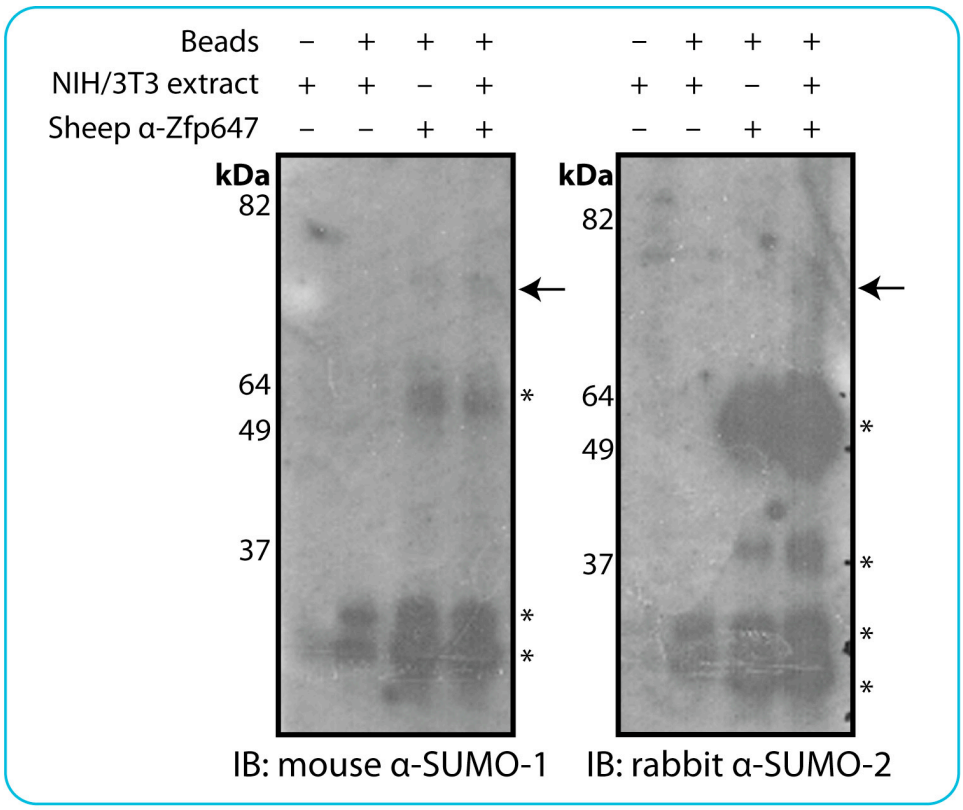


Figure 4.8. Immunoprecipitation of modified Zfp647 by sheep α-Zfp647. Western blot of immunoprecipitations of Zfp647 from NIH/3T3 whole cell extracts in the absence or presence of NEM, probed with the rabbit α-Zfp647 antibody. Asterisk indicates IgG bands; unfilled arrow indicates non-specific cytoplasmic band detected by the rabbit α-Zfp647 antibody.



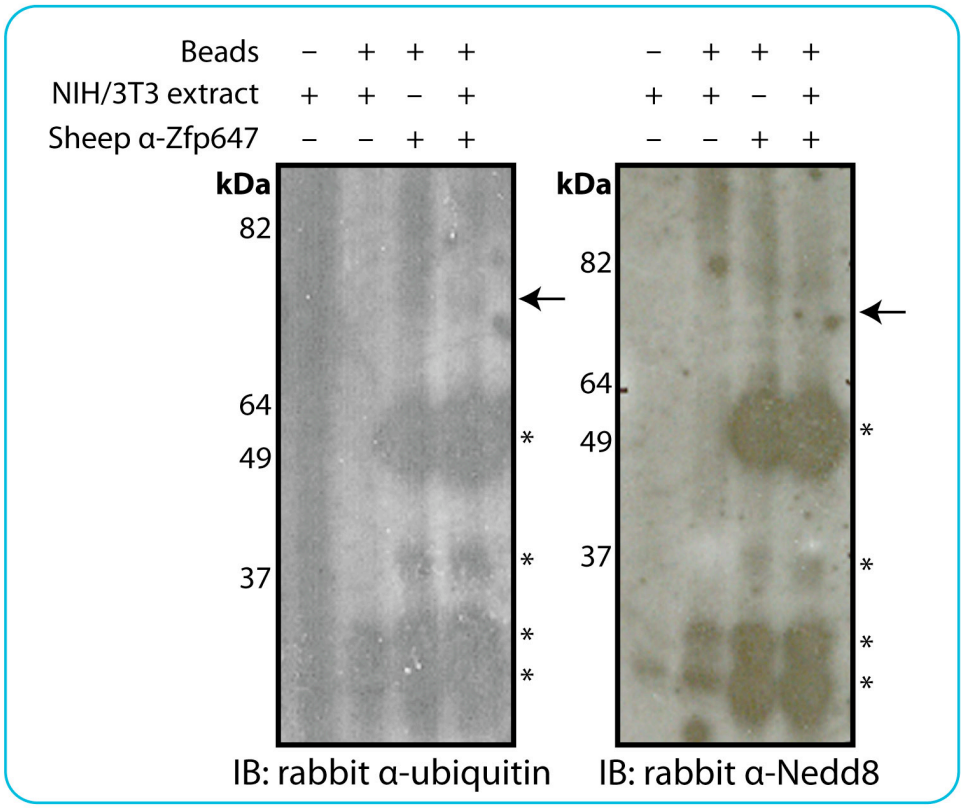


Figure 4.9. Immunoprecipitation of modified Zfp647 and detection with α-ubiquitin-like modifer antibodies. Western blots of NIH/3T3 whole cell extracts immunoprecipitated with sheep α-Zfp647 in the presence of NEM, probed with antibodies against SUMO-1, SUMO-2, ubiquitin and Nedd8. Arrows indicate the expected size of modified Zfp647; asterisks indicate IgG and non-specific bands.

ubiquitin-like modifiers, and the extra weight of the epitope tag would increase the size of the modified band, as shown with KAP-1 (Mascle *et al*, 2007). Thus this method could demonstrate if any of the modifiers was responsible for the higher molecular weight forms of Zfp647. I received plasmids encoding His_6-tagged ubiquitin (Dr Lee Finlan, MRC Human Genetics Unit), SUMO-1 and SUMO-2 (Professor Ron Hay, University of Dundee) and Nedd8 (Dr David Girdwood, MRC Human Genetics Unit). NIH/3T3 cells were transfected with these plasmids, and whole cell extracts of untransfected and transfected cells were western blotted with a mouse _-His antibody to check the expression of each of the proteins. The SUMO-1 plasmid from Ron Hay did not seem to express SUMO-1 (Figure 4.10A), and therefore another plasmid encoding this protein was obtained from Professor Anne Dejean (Institut Pasteur, Paris). When extracts were made from cells transfected with this plasmid, SUMO-1 was clearly expressed, seen at the bottom of the gel (Figure 4.10B). A band corresponding to SUMO-2 is seen at the bottom of the gel in the third lane of Figure 4.10A, and a dark smear at high molecular weight suggests that this protein is being incorporated into high molecular weight proteins. SUMO-2/-3 contains the four amino acid sumoylation consensus sequence ψKxD/E, and so can be incorporated into polySUMO chains, whereas SUMO-1 does not contain the consensus motif, thus the inclusion of SUMO-1 into chains stops their growth (Tatham *et al*, 2001). Free ubiquitin and Nedd8 were not seen within the cell extract (Figure 4.10A), but cells transfected with these proteins give a strong smear at high molecular weight, again suggesting that the protein is being expressed and attached to substrate proteins. Therefore cells transfected with these plasmids can successfully express the encoded modifier proteins.

Nuclear extracts of cells transfected with each of the His_6-tagged modifying proteins were analysed for any possible increase in the amount of the modified forms of Zfp647, or an increase in the molecular weight corresponding to the extra weight of the His_6 tag, which has a molecular weight of ~1kDa. The extracts were western blotted with sheep _-Zfp647, and with mouse _-tubulin (Figure 4.11). The level of unmodified Zfp647 is equal in all lanes, as is the level of the 85kDa modified Zfp647 band. There are two lower molecular weight modified bands present in the extracts

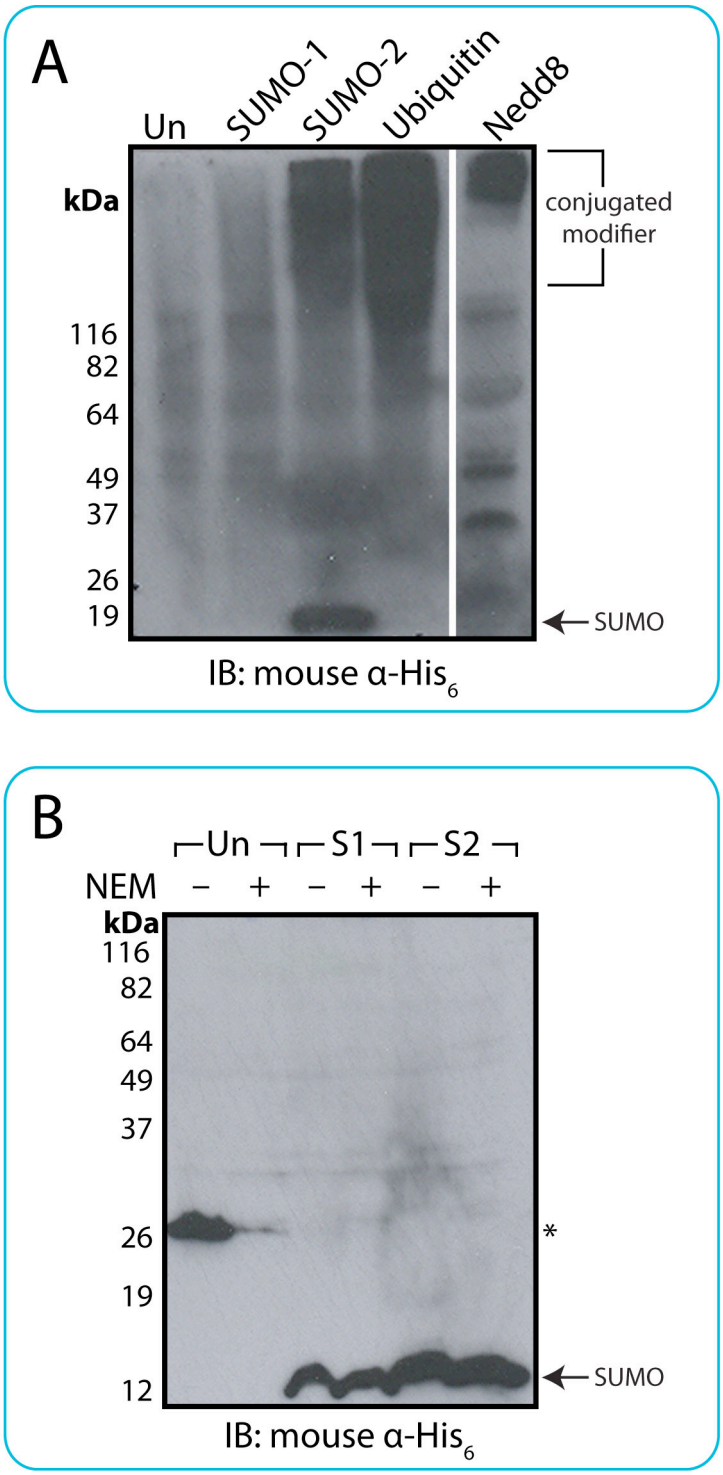


Figure 4.10. Expression of His_6-tagged ubiquitin-like modifiers in NIH/3T3 cells. A) Western blot of whole cell extracts made from cells transfected with SUMO-1, SUMO-2, ubiquitin, Nedd8 or untransfected cells (Un), probed with mouse α-His_6. The SUMO-1 plasmid from Prof Ron Hay was used in this experiment. All extracts were prepared in the presence of NEM. B) Western blot of whole cell extracts from untransfected NIH/3T3 cells, or cells transfected with His_6-tagged SUMO-1 or SUMO-2, probed with mouse α-His_6. The SUMO-1 plasmid from Prof Anne Dejean was used in this experiment. Extracts were made in the presence or absence of NEM. Asterisk indicates unknown band.

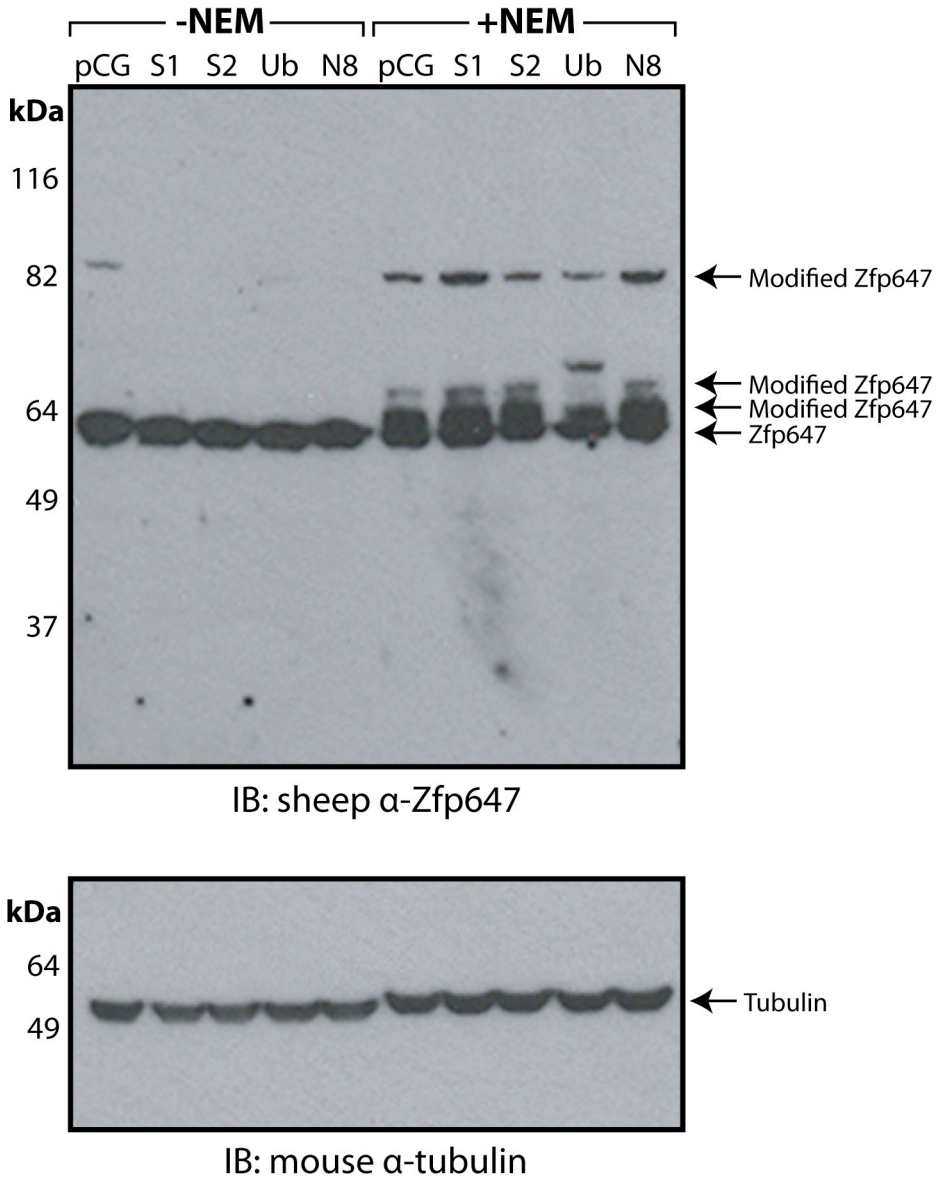


Figure 4.11. Effect of transfected ubiquitin-like modifiers on Zfp647 modification. Western blot of nuclear extracts made from NIH/3T3 cells transfected with the empty pCGT7 vector (pCG) or plasmids carrying His_6-tagged SUMO-1 (S1), SUMO-2 (S2), ubiquitin (Ub) or Nedd8 (N8). Extracts were blotted with sheep α-Zfp647 antibody (upper panel) or mouse α-tubulin (lower panel).

made with NEM, and in the cells transfected with the ubiquitin plasmid, the upper of these two bands appears to have increased in size. The other bands in this lane remain the same size, therefore this experiment suggests that this modification band may result from the mono-ubiquitylation of Zfp647. Interestingly, the 85kDa band can be seen at a low level in the –NEM extract prepared from cells transfected with the pCGT7 vector, but not in those transfected with each of the ubiquitin-like modifiers. It may be that the extra modifiers have an effect on the cell's protein turnover machinery, whereby a cell overloaded with the ubiquitin-like proteins causes the degradation of some, or maybe a large proportion, of the cell's ubiquitinated, sumoylated, neddylated, or otherwise modified proteins. It may therefore by interesting to analyse whether this turnover is also seen in cells inhibited with MG132, the proteasome inhibitor, which could possibly play a part in such a pathway.

4.3.4. Nickel Pull-Down of His_6-Tagged Modified Proteins

To try to verify whether Zfp647 is ubiquitylated, or possibly sumoylated or neddylated, I carried out a nickel pull-down, which can be used to demonstrate the modification of a substrate by one of the ubiquitin-like proteins (Rodriguez *et al*, 2001; Seeler *et al*, 2001). In this method, His_6-tagged modifiers are expressed in tissue culture cells, and become covalently attached to their acceptor proteins. The modified proteins can be purified due to the fact that the histidine residues comprising the tag bind with very high affinity to Ni^{2+}-NTA agarose. Proteins pulled down by the nickel agarose can then be probed for the presence of the suspected substrate.

I transfected NIH/3T3 cells with His_6-tagged SUMO-1, SUMO-2, ubiquitin or Nedd8, or the empty pcDNA3.1 vector. These cells were lysed in a buffer containing 6M guanidium hydrochloride, which completely denatures proteins, disrupting any non-covalent interactions both within the protein structure and between interacting proteins. Since ubiquitin-like proteins are covalently attached to their substrates,

these modifications remain as part of the substrate protein. The modifications cannot be removed from their substrates in this buffer, as the SENPs, DUBs and Nedd8 isopeptidases are also denatured by the guanidium hydrochloride treatment. The cell lysates were then incubated with Ni^{2+}-NTA agarose beads. These beads bind to the His_6-tag, even in its denatured state, and can therefore be used to pull down any factors covalently modified by a His_6-ubiquitin-like protein. The beads were washed, and the proteins associated with them were analysed by western blot. As a positive control, the pulled-down proteins were probed for the presence of RanGAP1, one of the major SUMO substrate protein in mammalian cells.

The inputs into the pull-downs and the proteins eluted from the nickel beads were run out by PAGE and western blotted with an _-RanGAP1 antibody (Figure 4.12A). RanGAP1 is present in all input lanes, and appears at a weight of approximately 85kDa. In the pull-down material, RanGAP1 is only detected in the lanes of the cells transfected with His_6-SUMO-1 and His_6-SUMO-2 (Figure 4.12A). This result is in agreement with published data that has shown that RanGAP1 can be modified by both SUMO-1 and SUMO-2 (Vertegaal *et al*, 2006). The RanGAP1 signal seen in the input lane is at approximately 85kDa, whereas the bands in the pull-down lanes are at approximately 95kDa, indicating that this is the modified form of RanGAP1.

The input fractions were also tested for the presence of Zfp647. Zfp647 was present in each of the transfected cell extracts (Figure 4.12B). However, when the pull-down material was probed for Zfp647, a band corresponding to this protein was found in each of the lanes (Figure 4.12B), suggesting that either Zfp647 is non-specifically pulled down by the nickel beads, or the antibody used to detect the protein cross-reacts with something present in the pulled down material. Therefore I repeated the western blots of the inputs and pull-down material with rabbit _-Zfp647 (Figure 4.13A). In the input lanes, the rabbit antibody primarily detects the non-specific cytoplasmic band at 40kDa, but also detects the Zfp647 band at 65kDa. Like the sheep _-Zfp647 antibody, however, the rabbit antibody gives a strong signal from every lane of the pull-down material.

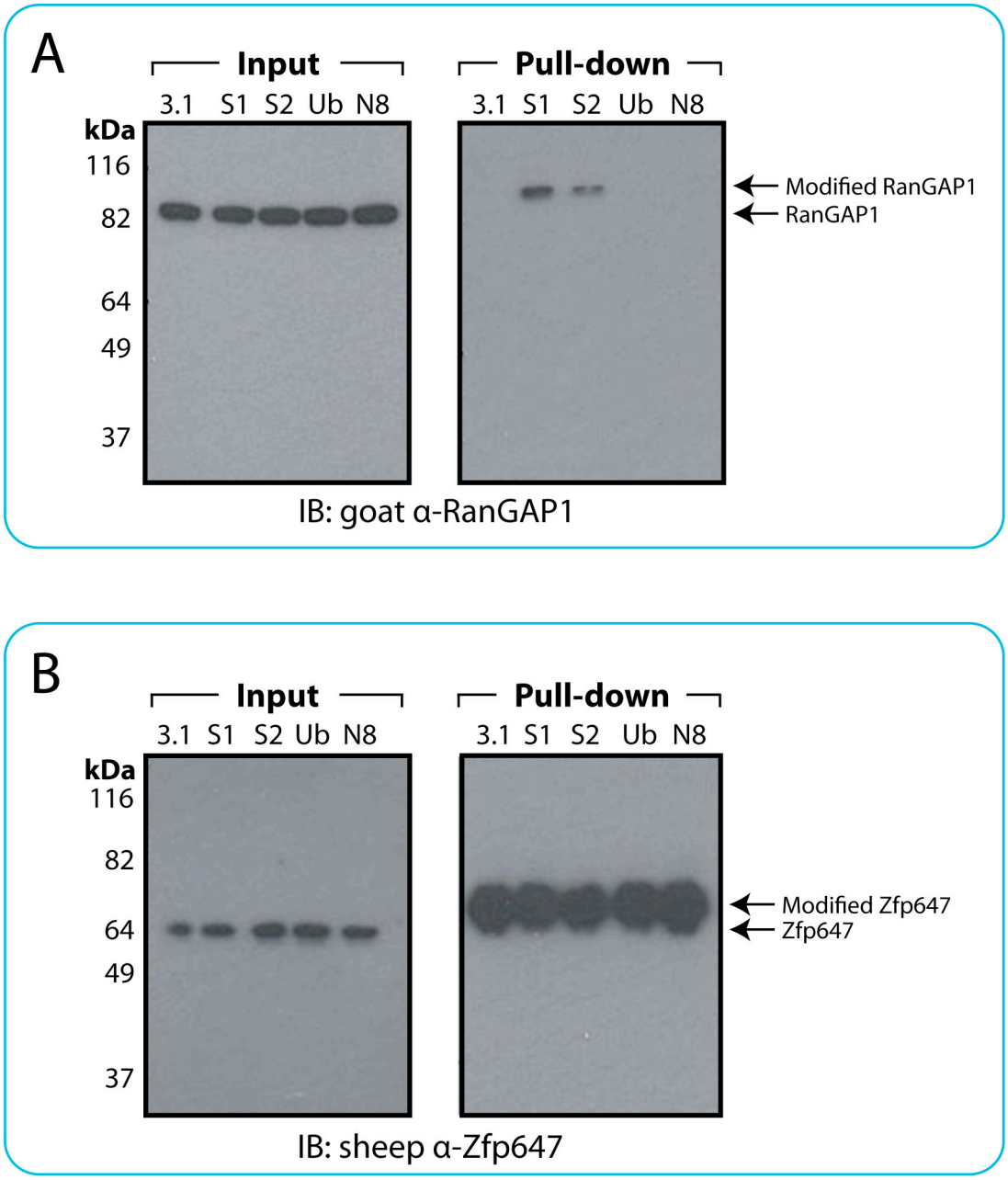


Figure 4.12. Nickel pull-down of proteins modified by His_6-tagged ubiquitin-like proteins. Western blots of nickel pull-downs carried out using extracts of NIH/3T3 cells transfected with pcDNA3.1 (3.1) or His_6-tagged SUMO-1 (S1), SUMO-2 (S2), ubiquitin (Ub) or Nedd8 (N8). Input (left-hand panels) and pull-down material (right hand panels) were western blotted with A) goat α-RanGAP1 or B) sheep α-Zfp647 antibodies.

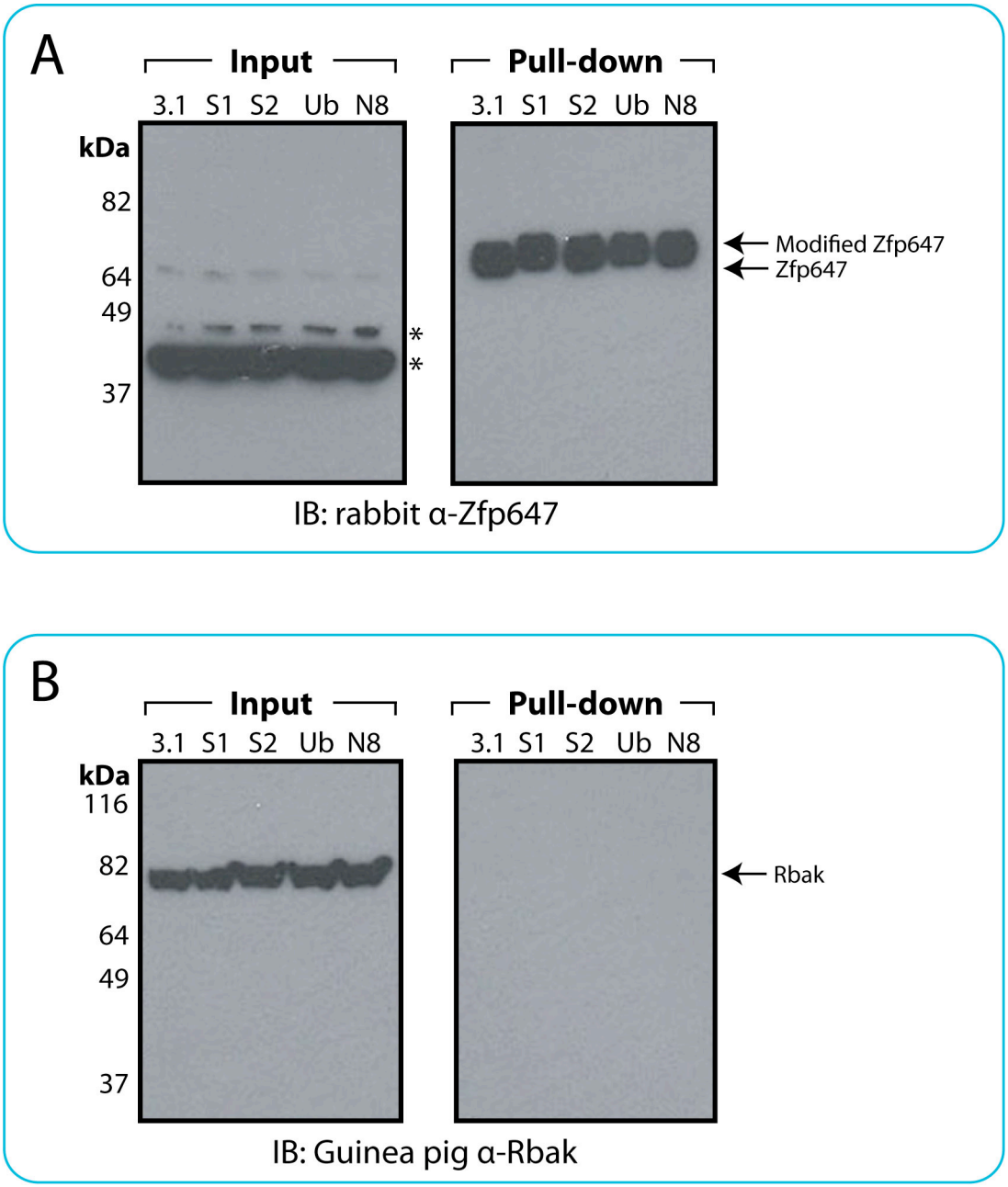


Figure 4.13. Nickel pull-down of proteins modifed by His_6-tagged ubiquitin-like proteins. Western blots of nickel pull-downs carried out using extracts of NIH/3T3 cells transfected with pcDNA3.1 (3.1) or His_6-tagged SUMO-1 (S1), SUMO-2 (S2), ubiquitin (Ub) or Nedd8 (N8). Input (left-hand panels) and pull-down material (right hand panels) were western blotted with A) rabbit α-Zfp647 or B) Guinea pig α-Rbak serum. Asterisks indicate non-specific bands.

Because of the presence of zinc fingers in Zfp647, it is possible that these motifs bind to the nickel agarose non-specifically. Zinc finger structures have been shown to bind to metals other than zinc (Makowski and Sunderman, 1992; Asmuß *et al*, 2000), and nickel, cadmium, cobalt and arsenic have been found to inhibit DNA repair pathways, possibly by alteration of zinc finger protein structure (Hartwig, 1998; Asmuß and Sunderman, 2000). I therefore tested whether Rbak behaved similarly to Zfp647 in this assay. I have been unable to detect sumoylated Rbak in NIH/3T3 cell extracts with Guinea pig _-Rbak unpurified serum. This may be because the antibody is unable to detect the modified forms of the protein, or may result from low levels of sumoylated Rbak in this cell-type. The input and pull-down fractions from the nickel pull-down were western blotted with Guinea pig _-Rbak unpurified serum (Figure 4.13B). Rbak is present in each of the input lanes (left panel), but no signal is detectable from the pull-down lanes (right panel). Thus Rbak does not non-specifically interact with nickel-agarose, unlike Zfp647, showing that this is not a trait common to all KRAB zinc finger proteins. The metal binding ability of zinc fingers has been shown to differ between proteins (Asmuß and Sunderman, 2000), therefore the zinc fingers of Zfp647 may be capable of binding to nickel, whereas Rbak's zinc fingers cannot. Rbak was not detected in the cells transfected with His_6-tagged SUMO, and so I was unable to demonstrate that Rbak is sumoylated in NIH/3T3 cells.

4.4. Modification of Zfp647 by Poly(ADP-Ribose)

4.4.1. Poly(ADP-Ribosyl)ation

My studies trying to determine what species Zfp647 is modified by appeared to suggest that this protein could be mono-ubiquitylated. However, mono-ubiquitylation only appeared to be responsible for one of the extra bands seen with NEM treatment, and I had found that two further modified bands could be seen (Figure 4.11).

Therefore I decided to test whether Zfp647 was subjected to another modification, poly(ADP-ribosyl)ation. Like the ubiquitin-like proteins, ADP-ribose can be conjugated to a substrate protein as a monomer by ADP-ribosyltransferases, or in chains by the poly(ADP-ribose) polymerases (PARPs) (Figure 4.14). ADP-ribosylation is associated with bacterial toxins, for example cholera toxin, which ADP-ribosylates G proteins within the intestines, leading to the characteristic fluid loss associated with this disease (Kaper *et al*, 1995). In Section 3.2.2, I showed that Zfp647 interacts with ARD1 in a yeast two-hybrid assay. ARD1 contains an ARF domain, which can stimulate ADP-ribosylation activity by proteins such as the cholera toxin, although ARD1 has not been directly shown to possess this activity. Poly(ADP-ribosyl)ation is commonly associated with DNA repair pathways. The founding, and probably most studied, member of the PARP family, PARP-1, interacts with and possibly poly(ADP-ribosyl)ates factors from both the base-excision repair and single-strand-break pathways upon formation of DNA nicks (Menissier de Murcia *et al*, 1997; Schreiber *et al*, 2006). Upon binding to nicked DNA, PARP-1 also poly(ADP-ribosyl)ates itself. Because of the high negative charge of the poly(ADP-ribose) mark, modified PARP-1 is released from the negatively charged DNA, and it has been suggested that this transient association of PARP-1 to damaged DNA regulates the DNA repair pathway (Smith, 2001). PARP-1 can also bind to PARP-2; PARP-2 has been shown to be required for efficient single-strand-break repair (Schreiber *et al*, 2002).

PARPs have also been implicated in the regulation of chromatin structure and transcription. PARP-1 can ADP(ribosyl)ate histone H1 (Section 1.1.3.5), which may lead to the removal of nucleosomes and the relaxation of chromatin structure (Ausió *et al*, 2001), although it has also been reported that PARP-1 can alter chromatin structure independent of its modification ability (Kim *et al*, 2004). PARP-1 is also known to have a stimulatory effect on transcription factors, for example NF-κB (Hassa *et al*, 2003) and the Mediator complex (Pavri *et al*, 2005). The insulator protein CTCF, which helps to prevent the preading of hetero- or euchromatic domains, is also subjected to poly(ADP-ribosyl)ation *in vivo* (Yu *et al*, 2004). Modification of CTCF does not disrupt its DNA-binding capability, as found with

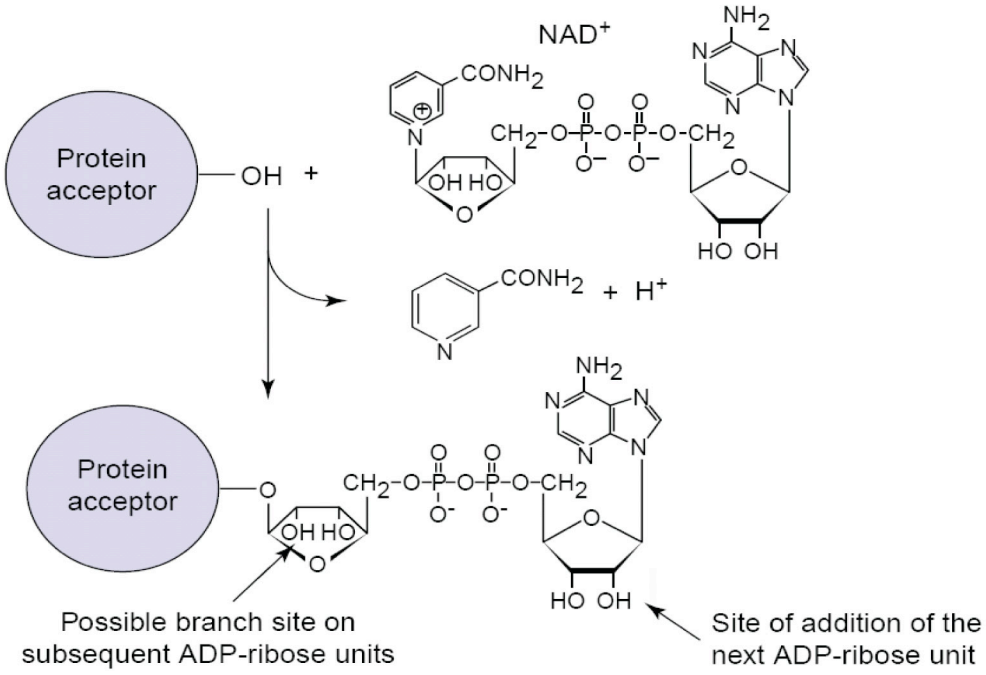


Figure 4.14. ADP-ribosylation of an acceptor protein. Using NAD^+ as a substrate, the acceptor protein is ADP-ribosylated on an OH group, for example at arginine, glutamic acid or aspartic acid residues. A protein that has been singly ADP-ribosylated may also be subjected to further ADP-ribosylation by PARP, generating a poly(ADP-ribosyl)ated protein. Figure taken from Smith, 2001.

PARP-1. Instead, poly(ADP-ribosyl)ation appears to be important for CTCF's insulator function, as cells treated with a PARP inhibitor, 3-aminobenzamide (3-AB), leads to a loss of insulator function at CTCF binding sites (Yu *et al*, 2004). Additionally, the nucleosome remodeller ISWI has been shown to be poly(ADP-ribosyl)ated in *Drosophila*, which inhibits the enzyme's function (Sala *et al*, 2008).

PARP-1 and PARP-2 have recently been linked to the regulation of cell differentiation by their effect on KAP-1 and HP1 (Quénet *et al*, 2008). It was shown that PARP-2 binds to KAP-1, HP1_ and HP1β; PARP-1 binds weakly to KAP-1 and HP1β; and that both PARPs were able to poly(ADP-ribosyl)ate HP1_ on its hinge domain. PARP-2 also co-localises with KAP-1 in primitive-endoderm-like (PrE) cells, generated by the differentiation of F9 EC cells, and the loss of PARP-2 leads to the impaired localisation of KAP-1 at the pericentric heterochromatin in these cells. Quénet *et al* also demonstrated that PARP-1 and PARP-2 may play a role in regulating the expression of mesoderm- and endoderm-specific genes through their interaction with and possible modification of KAP-1 and HP1 (Quénet *et al*, 2008). As it can regulate the behaviour of KAP-1 and HP1, I speculated that poly(ADP-ribosyl)ation may be involved in the regulation of other transcription factors.

4.4.2. Immunoprecipitation of Poly(ADP-Ribosyl)ated Proteins

To see whether Zfp647 is poly(ADP-ribosyl)ated, I carried out an immunoprecipitation using a mouse _-poly(ADP-ribose) antibody on NIH/3T3 nuclear extract, and then western blotted the immunoprecipitated material with the sheep _-Zfp647 antibody (Figure 4.15). As controls, I repeated the experiment substituting the same amount of IgG for the mouse _-poly(ADP-ribose) antibody. The experiment was carried out in the presence of NEM to preserve the modified forms of Zfp647, but a nuclear extract from cells not treated with NEM was run to show unmodified Zfp647. The higher molecular weight forms of Zfp647 can be seen in the input lane, and the lower of the two main modified forms also appears to be present in the immunoprecipitation lane, therefore this experiment suggested that

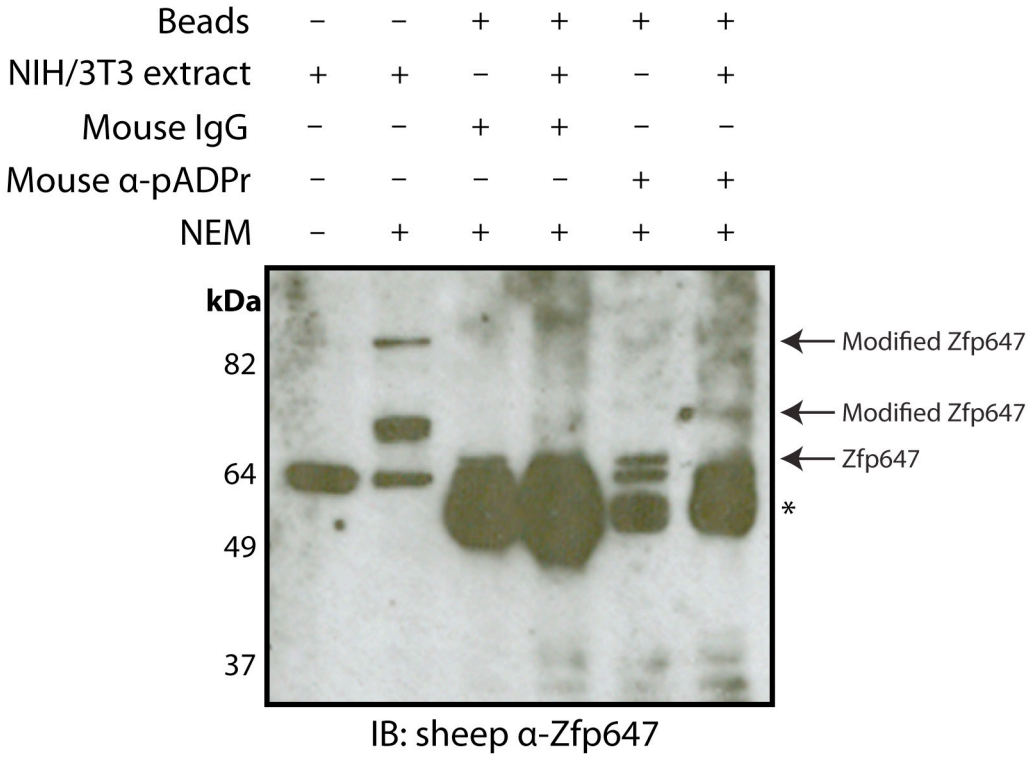


Figure 4.15. Detection of poly(ADP-ribosyl)ated Zfp647. Western blot of NIH/3T3 nuclear extract immunoprecipitated with either mouse α-poly(ADP-ribose) (pADPr) antibody or mouse IgG, probed with sheep α-Zfp647. Both immunoprecipitations were carried out in the presence of NEM. Asterisk indicates IgG bands.

Zfp647 may indeed be modified by poly(ADP-ribose). I also carried out the immunoprecipitation using rabbit _-Zfp647 antibody, and western blotted the immunoprecipitated material with the mouse _-poly(ADP-ribose) antibody; however, the western blots did not show any bands specific for Zfp647 (data not shown).

4.4.3. Effect of PARP Inhibition on Zfp647 Modification

To verify whether Zfp647 is poly(ADP-ribosyl)ated, I analysed the modification status of Zfp647 in cells in which PARP function was inhibited by treatment with 3-AB (Purnell and Whish, 1980). 3-AB, like other PARP inhibitors, functions by competing with free nicotinamide adenine dinucleotide (NAD^+), which is required as a precursor and substrate in the poly(ADP-ribosyl)ation reaction (Figure 4.14; Hassa and Hottiger, 2008). 3-AB is commonly used at concentrations of 2 or 3mM, typically for up to 24 hours (Kuo *et al*, 1996; Rajesh *et al*, 2006; Ding *et al*, 2008). I treated NIH/3T3 cells with 2mM 3-AB for 18 hours, and made whole cell extracts from these cells, as well as cells treated with the 3-AB solvent DMSO for the same time period. Extracts were made from cells both in the absence and presence of NEM, and were run out by PAGE. The gels were then western blotted with sheep _-Zfp647 and mouse _-tubulin antibodies (Figure 4.16). The western blot probed with the mouse _-tubulin antibody shows that the loading of each of these extracts was equivalent. In the western blot probed with sheep _-Zfp647, the 65kDa Zfp647 band is visible in each of the lanes. In the extracts treated with NEM, the modified forms of Zfp647 seen at approximately 70 and 80kDa are present in the cell extracts treated with both DMSO and 3-AB. In this experiment, it appeared that treatment of cells with 3-AB did not prevent the modification of Zfp647, raising doubts as to whether Zfp647 is poly(ADP-ribosyl)ated. However, as I have not included a positive control for the action of 3-AB, I cannot prove that the PARPs have been inhibited, therefore this result is inconclusive.

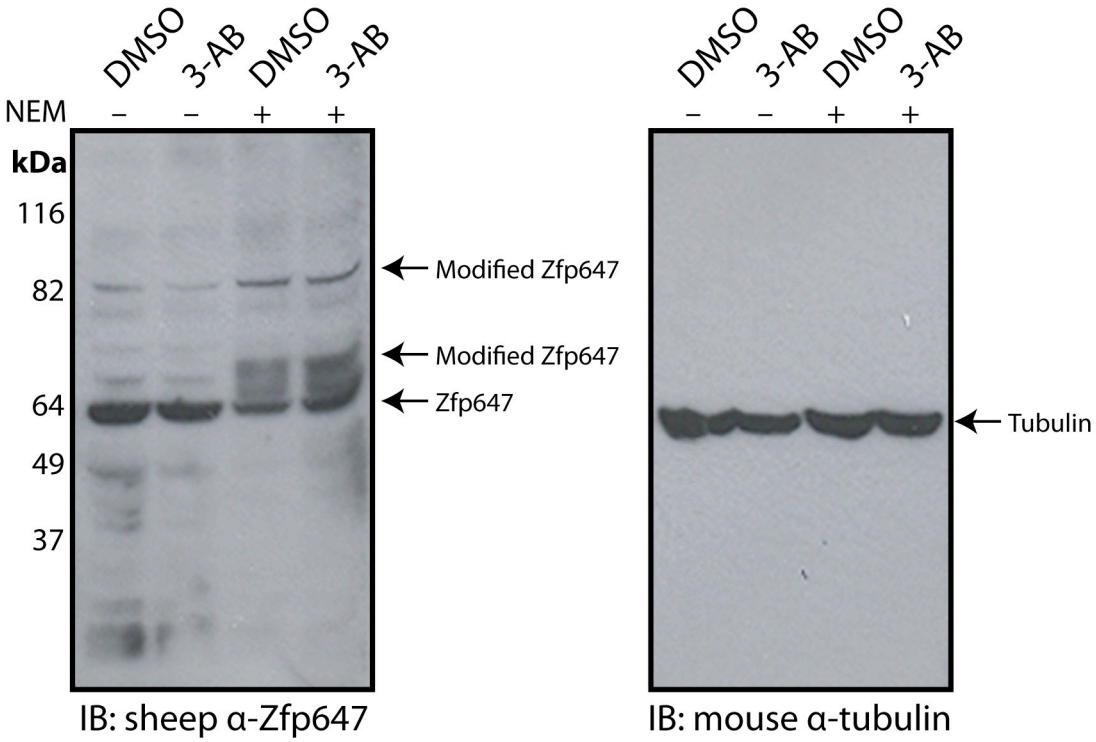


Figure 4.16. Effect of 3-aminobenzamide (3-AB) on Zfp647 modification. Western blots of nuclear extracts made from NIH/3T3 cells treated with 2mM 3-AB or DMSO for 18 hours, probed with either sheep α-Zfp647 or mouse α-tubulin.

4.5. Modification of Zfp647 in ES Cells

4.5.1. Relocalisation of Zfp647 in Differentiated ES Cells

The relocalization of Zfp647 in ES cells over the course of differentiation has been previously studied by Dr Heidi Sutherland, using OS25 cells (Section 1.4.2). In undifferentiated cells, Zfp647 shows a nuclear diffuse pattern of staining, but upon RA-induced differentiation, the protein relocalises to non-heterochromatic foci (Figure 1.9; Briers *et al*, 2009). The percentage of cells showing these KAKA foci is graphed in Figure 4.17. Zfp647 is first seen at non-heterochromatic foci after four days of differentiation: 2% of cells show this pattern of staining. After six days of RA differentiation, Zfp647 localises to non-heterochromatic foci in 44% of cells. In 38% of these cells, KAP-1 co-localises with Zfp647. The prevalence of the foci decreases after this point, and after ten days of differentiation, the Zfp647 foci are only seen in 7.5% of cells.

4.5.2. Modification of KAP-1 in ES Cells

It has been reported that each of the SUMO isoforms can be visualised at the PML bodies, resulting from the high concentration of sumoylated PML and other factors located at these structures (Boddy *et al*, 1996; Briers *et al*, 2009). As KAP-1 is also sumoylated, and can be visualised at structures adjacent to these SUMO-containing PML-NBs in differentiated ES cells, Dr Heidi Sutherland tested whether the modified forms of KAP-1 could be detected in ES cells, and also whether the modification of KAP-1 is different in differentiated cells compared with undifferentiated cells, as the localisation of this protein does. In whole cell extracts prepared with NEM, unmodified KAP-1 could not be seen in undifferentiated ES cells. However, in cells differentiated with 10 days of RA treatment, a band ~50kDa greater than endogenous KAP-1 could also be seen by western blot, indicating that KAP-1 is modified, probably by sumoylation, in differentiated, but not undifferentiated, ES cells (Briers *et al*, 2009).

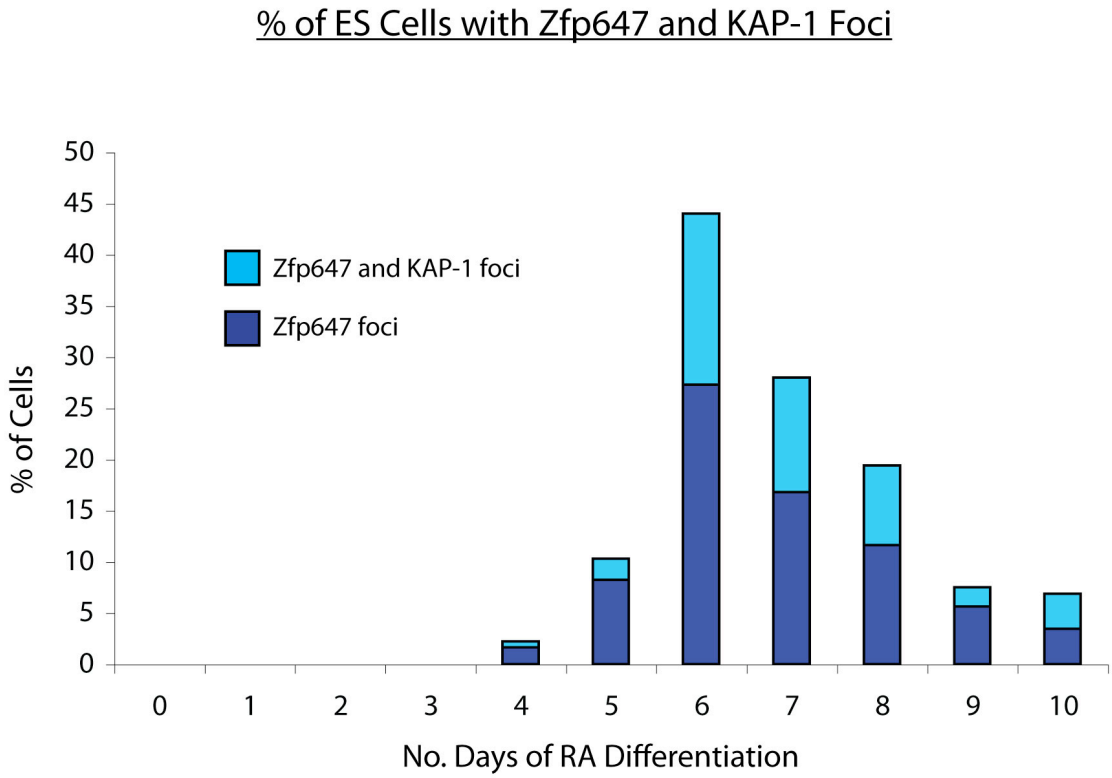


Figure 4.17. Prevalence of KAKA foci in ES cells over the course of differentiation. Graph shows the percentage of OS25 cells displaying either Zfp647 foci (dark blue) or Zfp647 and KAP-1 foci (light blue) over a 10-day period of retinoic acid (RA) treatment. All data was generated by Dr Heidi Sutherland.

4.5.3. Detection of Modified Forms of Zfp647 in ES Cells

Because the modification of transcription factors has been linked with their ability to co-localise with other proteins and to nuclear structures (Zhong *et al*, 2000; Quénet *et al*, 2008), and the results obtained by Dr Heidi Sutherland suggested that KAP-1 modification is dynamic over the course of ES cell differentiation, I decided to see whether the modification status of Zfp647 also changes as ES cells differentiate. I chose to use OS25 cells, as these cells had been used to analyse the localisation of Zfp647. OS25 cells were first generated to allow the differentiation of pluripotent cells to neuroepithelial or oligodendrocyte precursor cells (Billon *et al*, 2002). Although RA induces the cells to differentiate down a neural pathway, I did not positively select for either neuroepithelial or oligodendrocyte precursor cells. OS25 cells have a *hygromycin-thymidine kinase* construct cloned into the *Oct4* locus. To differentiate the cells, I grew them in the absence of leukaemia inhibitory factor (LIF) and in the presence of RA, and also selected against undifferentiated, Oct4-expressing, cells by treating them with gancyclovir.

Whole cell extracts were made from OS25 cells before the differentiation process was initiated, and then 1, 3, 5, and 7 days after the LIF was first removed from the growth media. Extracts were made in the absence and presence of NEM, run out by PAGE and western blotted with sheep _-Zfp647 (Figure 4.18). The relative protein levels of these extracts was assayed by western blotting the extracts for histone H3. Zfp647 is present in each of the lanes, showing that the expression of this protein does not change during OS25 differentiation. This corroborates Dr Heidi Sutherland's immunofluorescence data, as Zfp647 could be seen throughout differentiation, and it was the localization of the Zfp647 signal that changed. In the undifferentiated cell extract, an extra band is present above the 65kDa Zfp647 band. After 1 and 3 days of differentiation, the extra band has almost disappeared, but reappears in the day 5 and 7 extracts. It therefore appears that the modification of Zfp647 alters over the period of differentiation: it is lost at the beginning of the differentiation process, but reappears as the cells become differentiated. Interestingly, the re-emergence of the 70kDa modified band coincides with the re-

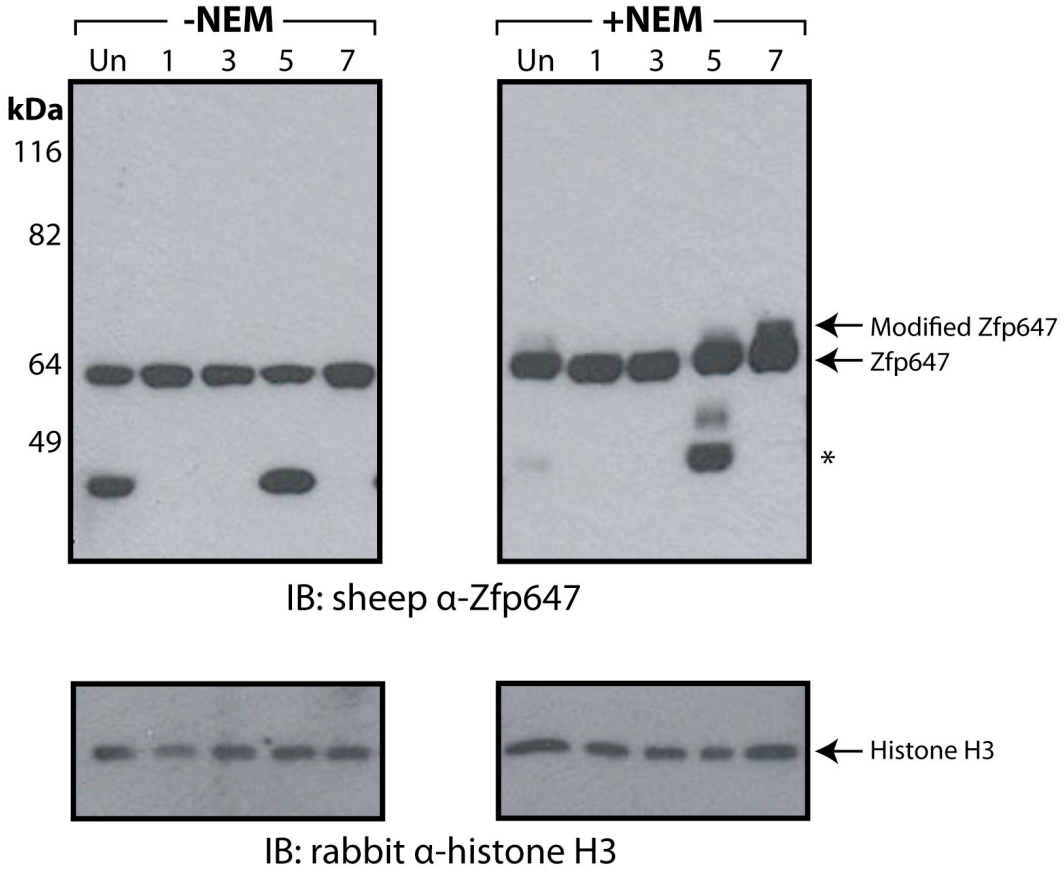


Figure 4.18. Zfp647 modification throughout ES cell differentiation. Western blots of whole cell extracts made from undifferentiated OS25 cells (Un) or OS25 cells differentiated with retinoic acid (RA) for 1, 3, 5 or 7 days. Extracts were detected with sheep α-Zfp647 or rabbit α-histone H3. Asterisk indicates possible Zfp647 degradation product or non-specific band.

localisation of Zfp647 to the KAKA foci. It may be possible that the modification of Zfp647 is in some way linked to the protein's localisation within the nucleus.

4.6. Discussion

4.6.1. Zfp647 is Post-Translationally Modified

KAP-1 has been shown to be sumoylated (Lee *et al*, 2007b; Mascle *et al*, 2007) and this modification is required for its interaction with two of the effector proteins in the KRAB-mediated transcriptional repression machinery, Mi-2_ and SETDB1 (Ivanov *et al*, 2007). Because KAP-1 functions as a SUMO E3 ligase, it may be possible that this protein can direct the modification of other proteins it interacts with, including the KRAB zinc finger proteins. Alternatively, the RING domain of KAP-1, or any of the RING domain proteins pulled out of my yeast two-hybrid screen, could facilitate the ubiquitylation (or sumoylation) of Zfp647. Therefore I tested whether Zfp647 is post-translationally modified *in vivo*. One of the most common reasons for protein modification is poly-ubiquitylation, which typically leads to a protein's degradation by the proteasome. As Zfp647 interacts with RING domain proteins that could function as ubiquitin E3 ligases, the amount of cellular Zfp647 may be regulated by its poly-ubiquitylation catalysed by these proteins. I therefore tested to see whether inhibition of the proteasome using the chemical MG132 had an effect on Zfp647. If Zfp647 is targeted for destruction, I would have expected to see an accumulation of modified forms of Zfp647 as they could no longer be removed from the cell. However, MG132 had no effect on the pattern of Zfp647 staining seen by western blot (Figure 4.3), therefore it is unlikely that Zfp647's association with RING domain proteins leads to its poly-ubiquitylation and subsequent degradation.

I have found that modified forms of Zfp647 are detected in the presence of NEM, a cysteine protease inhibitor, in NIH/3T3, PCC4 EC and ES cells. Treatment of cells with NEM leads to the appearance of typically two extra Zfp647 bands, at either 70 or 75kDa, and 85kDa. The size of the lower band is not as constant as that of the 85kDa species (compare Figure 4.4 with Figure 4.11), thus it may be that there are at least three different modified forms of Zfp647, or these forms may not behave consistently on PAGE gels. Another band, seen in Figures 4.6 and 4.18, only appears to be 1-2kDa larger than the unmodified form of Zfp647. Since NEM is a general inhibitor of proteins that depend on cysteine residues in their active sites, it could prevent the action of many enzymes in the cell. However, given the indications from the yeast two-hybrid screen that Zfp647 may interact with the ubiquitin (or ubiquitin-like protein) machinery, I decided to try and see whether Zfp647 is modified by one of the three major ubiquitin-like proteins: ubiquitin, SUMO or Nedd8. The size of two of the extra Zfp647 bands, at approximately 10 and 20kDa above the main Zfp647 band, correlate with the size of one and two covalently attached ~10kDa ubiquitin-like modifiers. Despite being able to successfully immunoprecipitate the modified forms of Zfp647, probing the immunoprecipitated material with antibodies specific for the ubiquitin-like modifiers did not confirm that any of these proteins was covalently attached to Zfp647 (Figure 4.9). However, given that there was barely any signal from each of the input lanes in these immunoprecipitations, the antibodies used may not be particularly good at detecting the endogenous modifier proteins. This problem, coupled with the seemingly small pool of modified Zfp647 immunoprecipitated by the sheep _-Zfp647 antibody (Figure 4.8) may have lead to the modification of Zfp647 by one of the ubiquitin-like proteins going undetected. In the future, this experiment could be repeated using cells transfected with soluble, tagged Zfp647, which may increase the amount of modified Zfp647 that is present, a step that is often required to visualise proteins covalently linked to the ubiquitin-like modifiers (Mascle *et al*, 2007).

4.6.2. Zfp647 is Possibly Mono-Ubiquitylated

Transfection of tissue culture cells with His_6-tagged ubiquitin-like modifiers has suggested that Zfp647 could be mono-ubiquitylated. One of the modified Zfp647 bands increased in size in the presence of His_6-tagged ubiquitin (Figure 4.11). This increase in size could be due to the extra molecular weight of the modifier's tag, which, consisting of six histidine residues, should add 1-2kDa to the protein's molecular weight. The exogenous ubiquitin protein did not affect any of the other NEM-dependent bands. Interestingly, the band that shifted in size was already 1-2kDa greater than the size of endogenous Zfp647, which may suggest that only Zfp647 that is already modified can be ubiquitylated. The modification of Zfp647 by tagged ubiquitin could be further tested by transfecting the cells with ubiqutin carrying a larger epitope, for example GFP, which would demonstrate more clearly the change in size caused by its ubiquitylation. I attempted to verify the mono-ubiquitylation of Zfp647 by using the nickel pull-down method, but Zfp647 non-specifically precipitates with the nickel beads used, thus this experiment was not successful. Unlike Zfp647, Rbak did not non-specifically interact with the beads, suggesting that this property is not shared by all KRAB zinc finger proteins. In western blots of cell extracts treated with NEM, I was unable to detect any modified forms of Rbak, and the antibody did not give any signal from the nickel pull-down material. Thus either this antibody is unable to detect modified forms of Rbak, as the modification could interfere with the antibody's binding to its epitope; the amount of modified Rbak present in NIH/3T3 cells is too small to be detected; or Rbak is not sumoylated in NIH/3T3 cells. Transfection of cells with soluble, tagged Rbak and the ubiquitin-like modifiers and analysis by western blot with antibodies specific for the epitope tags could therefore be used to confirm whether or not Rbak is modified in these cells.

There are two main approaches that I would take in the future to see whether Zfp647 is post-translationally modified by one of the ubiquitin-like proteins. First, TNT-expressed Zfp647 could be subjected to *in vitro* ubiquitylation, sumoylation or neddylation assays. Zfp647 can be successfully expressed using the TNT system, and

these assays do not require endogenous protein. Since the putative E2 (in the case of ubiquitin) or E3 ligases are unknown, the assay could be supplemented with NIH/3T3 cell extract to allow efficient modification of Zfp647 (Mouchantaf *et al*, 2006). If any of these experiments showed that Zfp647 was modified by ubiquitin, SUMO or Nedd8, the region of the protein that harbours the modified lysine residue(s) could be narrowed down, and the target lysine(s) identified by mutation of the plasmid that Zfp647 was expressed from, an advantage that this system has over analysis of the endogenous protein. The other approach that could be taken would be to immunoprecipitate modified Zfp647 and subject the immunoprecipitated material to mass spectrometry, which can be used to identify modification by the ubiquitin-like proteins, as well as modifications such as acetylation and phosphorylation (Witze *et al*, 2007). To increase the amount of material to be analysed, I could over-express Zfp647 in the cell, as I have found that His_6-tagged Zfp647 is modified in the presence of NEM (data not shown).

I have also attempted to see whether Zfp647 is modified by another type of modification, poly(ADP-ribose). The PARPs are known to interact with KAP-1 and HP1, and may poly(ADP-ribosyl)ate these proteins *in vivo* (Quénet *et al*, 2008). Poly(ADP-ribosyl)ation can result in varying increases in the size of the substrate protein: HP1 poly(ADP-ribosyl)ation leads to an increase in size of <5kDa (Quénet *et al*, 2008), whereas poly(ADP-ribosyl)ation of CTCF was visualised as an increase in size of ~40kDa (Yu *et al*, 2004), therefore any of the modified forms of Zfp647 could be candidates for this modification. While NEM has been reported to be inhibitory to the ADP-ribosylation process, as it can covalently modify potential cysteine acceptor sites (Hoshino *et al*, 1990; van Heyningen and Baxty, 1997), it may be possible that NEM inhibits the removal of the poly(ADP-ribose) from Zfp647 by altering the activity of poly(ADP-ribose) glycohydrolase (PARG), or one of the other enzymes that reverse poly(ADP-ribosyl)ation (Davidovic *et al*, 2001). I carried out an immunoprecipitation of modified Zfp647 from tissue culture cells, and probed the immunoprecipitated material with an antibody specific for poly(ADP-ribose) (Figure 4.15). This suggested that the lower of the two modified bands may indeed by poly(ADP-ribosyl)ated Zfp647, and so I further tested this by inhibiting PARP

activity with 3-AB. Analysis of Zfp647 in cells treated with this chemical showed that there was no effect on Zfp647 modification. However, this experiment did not include a positive control to prove that the inhibitor was in fact preventing poly(ADP-ribosyl)ation by the PARPs. The experiment therefore needs to be repeated with an assay for poly(ADP-ribosyl)ation, possibly by detecting the forms of proteins known to be poly(ADP-ribosyl)ated *in vivo*, such as PARP-1 (Smith, 2001) or NF-κB (Hassa *et al*, 2003). Alternatively, an *in vitro* poly(ADP-ribosyl)ation assay could be employed (Yu *et al*, 2004), which would show if Zfp647 can be modified in this manner. Co-immunoprecipitations between either endogenous or over-expressed Zfp647 and the PARP enzymes could also be carried out, which would strengthen the argument that Zfp647 is poly(ADP-ribosylated).

4.6.3. KAP-1 Does Not Influence Zfp647 Modification

I have also tested whether KAP-1 itself influences Zfp647 modification. KAP-1's PHD and bromodomains function as a SUMO E3 ligase, catalysing KAP-1 sumoylation by Ubc9, and this protein also possesses a RING domain, which may act as a ubiquitin, SUMO or Nedd8 E3 ligase. Zfp647 modification was tested in EC cells in which KAP-1 was downregulated through an integrated RNAi construct (Figures 4.5 and 4.6; Wolf *et al*, 2008). These results showed that Zfp647 modification was not affected by the loss of KAP-1, but the amount of Zfp647 visible was slightly reduced in the PCC4 111 cells. Therefore it appears unlikely that KAP-1 modifies, or aids in the modification of, Zfp647. The expression of another KRAB zinc finger protein, Zfp809, has also been found to be reduced in the PCC4 111 cell line when compared with wild-type cells (Dr Daniel Wolf, personal communication). KAP-1 may instead be involved in regulating the expression or stability of KRAB zinc finger proteins. KAP-1 has been shown to bind to chromatin at the sites of the KRAB zinc finger protein genes, and H3K9 methylation is also found at these loci (O'Geen *et al*, 2007). It is therefore difficult to picture how the loss of KAP-1 causes a decrease in transcription of these genes, unless KAP-1 binding does not repress KRAB zinc finger protein gene expression, as suggested by

Vogel *et al* (Vogel *et al*, 2007). The effect of KAP-1 on transcription of Zfp647 expression could be tested relatively easily by qRT-PCR of the *Zfp647* gene in the wild-type and 111 PCC4 cells. Alternatively, KAP-1 may in some way stabilise KRAB zinc finger proteins. Binding to KAP-1, and possibly other KRAB-mediated transcriptional repressor proteins, may prevent KRAB zinc finger protein turnover, similar to the stabilisation of Ezh2 by the PRC proteins Suz12 (Pasini *et al*, 2004) and Eed (Montgomery *et al*, 2005). The loss of Zfp647 protein stability could also be suggested by qRT-PCR data, as if the *Zfp647* mRNA level does not change, but the protein level decreases, KAP-1 may be important for Zfp647 stability.

4.6.4. Zfp647 Modification is Dynamic During ES Cell Differentiation

We have previously found that Zfp647 localisation is dynamic over the course of ES cell differentiation (Briers *et al*, 2009). The modification of certain proteins, including transcription factors, has been shown to impact upon their localisation, particularly to the PML bodies. For example, sumoylated orphan nuclear receptor LRH-1 is only found at the PML-NBs, and release from the PML-NBs occurs at the same time as the protein's desumoylation (Chalkiadiki and Talianidis, 2005). Therefore, I investigated whether the modification of Zfp647 is also dynamic in ES cells. I found that the larger, 85kDa, modified form of Zfp647 was only present at low levels in ES cells. The smallest modified form of Zfp647 was not always seen (Figure 4.18). This form was present at low levels in undifferentiated ES cells, but disappeared upon treatment with RA, only to reappear 3-5 days later. The modification of Zfp647 correlates with the protein's localisation, as the KAKA foci are first seen in OS25 cells after 4 days of differentiation (Figure 4.17). It is therefore tempting to speculate that the protein's modification is linked to its localisation, and that modification may be required for recruitment to the KAKA foci. I do not believe that this modified form of the protein is sumoylated, as difference in molecular weight to the endogenous protein is too small. It may therefore be possible that another type of modification is involved in the recruitment of proteins to the KAKA foci. If the species of modification was successfully identified using one of the

methods outlined above, it may then be possible to chemically inhibit the modification of Zfp647, and visualize whether this impairs the recruitment of Zfp647 to the KAKA foci in differentiated ES cells.

An extra band is visible in Figure 4.18, at approximately 40kDa. This may be a Zfp647 degradation product, as the 65kDa band appears to be slightly weaker in strength in those lanes where this extra band is present (undifferentiated and day 5 lanes). This may therefore be a product of Zfp647 turnover; that controlled Zfp647 degradation is occurring at these timepoints. Interestingly, this band appears when the modified Zfp647 species is present. One possible reason for this turnover is that in order to change the localisation and/or modification status of Zfp647, the pre-existing quotient of the protein is degraded, breaking down any Zfp647 structures, allowing the newly produced, potentially modified, Zfp647 to become concentrated at the KAKA foci.

Chapter 5: Discussion

5.1. Summary

The KRAB zinc finger proteins are a very large family of transcriptional repressors found in mammalian genomes. Although the precise functions of most of these proteins are yet to be identified, it is thought that they can play important roles in such divergent processes as the regulation of DNA methylation (Li *et al*, 2008), sexual dimorphism (Krebs *et al*, 2003), correct positioning of the body axis (García-García *et al*, 2008) and the prevention of retroviral replication (Wolf and Goff, 2009). As the vast majority of the KRAB domain-containing proteins are associated with a zinc finger domain, it is thought that the main way in which KRAB zinc finger proteins exert an effect is through binding to a particular gene target and recruiting other factors to supress its transcription. These gene targets may be unique to one KRAB zinc finger protein, or could possibly be shared by closely related proteins encoded in a gene cluster. While the search for gene targets of the KRAB zinc finger proteins has not yielded many results, investgations into the protein-protein interactions of these factors has uncovered possible functions for a number of proteins, for example ZBRK1 (Zheng *et al*, 2000), Zkscan17 (Nielsen *et al*, 2004) and NRIF (Casademunt *et al*, 1999).

Like the number of known endogenous gene targets, the number of good antibodies that have been raised against the KRAB zinc finger proteins is also low. This may be due to high levels of homology found between certain proteins, leading to difficulties in raising antibodies with a high level of specificity, or alternatively may result from the KRAB zinc finger proteins' low profile within the biological field. Therefore most studies that have analysed the localisation of these proteins within the cell have depended on overexpression of tagged KRAB zinc finger protein constructs. We have shown that the localisation of GFP-Zfp647 in fact differs to that of the endogenous protein, as we have been able to raise a specific antibody against Zfp647 (Briers *et al*, 2009). This may also be true for other over-expressed KRAB zinc finger protein fusion proteins. We have found that endogenous Zfp647, and NT2, another murine KRAB zinc finger protein (Tanaka *et al*, 2002), localise to distinct non-heterochromatic foci in undifferentiated ES cells, alongside KAP-1, HP1

and SETDB1 (Briers *et al*, 2009). A very similar pattern of localisation has been reported for another KRAB zinc finger protein, PAROT (Fleischer *et al*, 2006). Two KRAB zinc finger proteins, KRAZ1 and KRAZ2, have been shown to be targeted to the centromeric heterochromatin through their interaction with KAP-1, and it is therefore postulated that KRAZ1 and KRAZ2 repress transcription of their targets by association with these domains of silent chromatin. Zfp647 and NT2 may silence their gene targets in a similar manner, by the recruitment of these genes to another transcriptionally silent environment, the KAKA foci. Because of their proximity to the PML bodies, it may instead be that these foci are associated with PML-NB function. Most relevant to the behaviour of transcription factors, the KAKA foci may, like the PML-NBs, be a site of storage of the KRAB zinc finger proteins, if they are not required at a particular timepoint or cell cycle stage. The formation of the foci may play a role in, or depend upon, the modification status of KAP-1 (Lee *et al*, 2007b; Mascle *et al*, 2007) or the KRAB zinc finger proteins (Vertegaal *et al*, 2006). I have found that the KAKA foci are not present in NIH/3T3 cells, whereas they are clearly visible in pMEFs. I showed that the foci in pMEFs also twinned with PML bodies, as has been shown in the ES cells (Briers *et al*, 2009). Therefore it is possible that the KAKA foci are cell-type specific: certain cell-specific factors may be required to induce the formation of these foci. Alternatively, they may be protein-specific: I did not see KAKA foci in NIH/3T3 cells with antibodies against Zfp647, Zfp37 and Zfp90; however, other proteins may co-localise with KAP-1 in these cells. The two cell types that we have seen the KAKA foci in (ES cells and pMEFs) suggests that they may only be present in primary or untransformed cell lines.

Chapter 3 details the work I have carried out to identify protein-protein interactions that Zfp647 can make. The yeast two-hybrid results that I have shown are not false positives demonstrate that Zfp647 is capable of self-interaction, and interacting with a number of other KRAB zinc finger proteins. My *in vitro* co-immunoprecipitation results suggested that the two tagged Zfp647 moieties specifically interact, although the signal from the interaction was fairly weak. Homo-oligomerisation has been demonstrated for one other KRAB zinc finger protein, ZBRK1, which, like Zfp647, interacted with itself through its C-terminus. Tan *et al*

showed that ZBRK1 homo-tetramerises, and that this interaction enhances ZBRK1's transcriptional repression activity (Tan *et al*, 2004b). Furthermore, the repressive effect exerted by the C-terminus was shown to be independent of ZBRK1's DNA-binding ability, thus the authors argued that higher-order structures of ZBRK1 may form on its DNA targets, which could aid in the repression of these sequences by recruitment of more silencing proteins. The finding that Zfp647 can self-interact suggests that other KRAB zinc finger proteins may function in this manner, and the KAKA foci we have observed may result from these higher-order structures. The repressive effect caused by the oligomerisation of ZBRK1 is independent of the protein's DNA binding ability. Therefore, it may be the case that various KRAB zinc finger proteins, interacting through their zinc finger domains, can aid in the repression of one protein's gene targets, without compromising the specificity of the sequences that are being silenced.

I have found that Zfp647 may interact with two jumonji domain-containing proteins, Jarid1b and Jmjd1c using two separate methods, although I have not been able to verify either of these interactions through independent means. These potential interactors are of interest because it has been shown that another KRAB zinc finger protein, Zkscan17, can interact with JARID2 (Mysliwiec *et al*, 2007). Jarid1b is a histone H3K4 demethylase (Seward *et al*, 2007; Yamane *et al*, 2007), and would thus aid in transcriptional silencing. Whether Jmjd1c possesses any demethylase activity has yet to be established. It has been found that the tethering of KAP-1 to a chromatinised DNA sequence causes a decrease in H3K4 methylation and H3K9 acetylation, and an increase in H3K9, H3K36 and H4K20 methylation (Sripathy *et al*, 2006). Thus far, HP1, SETDB1 and Mi-2_ have been shown to function in the KRAB-mediated repression pathway, accounting for the decreases in histone acetylation and increase in H3K9 methylation at KRAB zinc finger protein binding sites. It may therefore be possible that the jumonji domain-containing proteins play a role in this pathway by removing pre-existing methyl marks on H3K4, as histone demethylase function has only been assigned to the jumonji proteins and LSD1 to date (Cloos *et al*, 2008).

Work that I have presented in Chapter 4 demonstrates that Zfp647 is post-translationally modified when treated with the cysteine protease inhibitor NEM. The most prominent extra Zfp647 bands appear at <5, 10 and 20kDa above the unmodified form of the protein when analysed by western blot. I have found evidence to suggest that Zfp647 may be mono-ubiquitylated, as transfection of tissue culture cells with a His_6-tagged ubiquitin construct causes the increase in size of one of these modified bands, although I have been unable to verify this result by other means. Out of the modifications that I was testing, I previously thought that the protein most likely to give a positive result was SUMO, as two KRAB zinc finger proteins have been shown to be sumoylated *in vitro* (Vertegaal *et al*, 2006), and KAP-1 acts as both a sumoylation substrate and its own E3 ligase (Lee *et al*, 2007b; Mascle *et al*, 2007; Ivanov *et al*, 2007). However, transcription factors such as c-myc, Tat and CIITA have been shown to be mono-ubiquitylated, which corresponds with an increase in their transcriptional activation ability (Bres *et al*, 2003; Greer *et al*, 2003; Kim *et al*, 2003; von der Lehr *et al*, 2003). Ubiquitylation of Zfp647 may also affect this protein's influence on transcription, possibly by inhibiting its interactions with other proteins. The co-immunoprecipitations shown in Figure 3.14 were carried out either in the presence or absence of NEM. Most of the bands that are seen in the co-immunoprecipitation lanes do not differ between these two lanes, however some, including the Jarid1b band, are specific to one lane. In the case of Jarid1b, this band is only seen in cells not treated with NEM. Therefore it may be that modification of Zfp647 prevents its interaction with this protein.

Finally, I have found evidence that the modification of Zfp647 is dynamic throughout ES cell differentiation. As we have previously shown that Zfp647 localisation also changes during differentiation, it may be possible that the two processes are linked, whereby modification is required for Zfp647 localisation to the KAKA foci, or modification actually occurs at these sites. The former of these two scenarios seems more likely, as Zfp647 modification occurs in NIH/3T3 cells, but Zfp647 localization to KAKA foci is not evident.

5.2. Future Perspectives

There are a number of experiments that immediately follow on from the work that I have presented in this thesis, some of which I have not been able to carry out in the time available. Firstly, to prove that Zfp647 is capable of interacting with the proteins identified in the yeast two-hybrid screen, these interactions should be tested in a second, independent system. While I have had difficulty expressing Zfp647 in a soluble form, many of these other proteins may more soluble, allowing co-immunoprecipitation experiments using endogenous Zfp647 and over-expressed, tagged interactor protein. The interactions that I would be most interested in initially are those of Zfp647 with itself, other KRAB zinc finger proteins, and the jumonji domain-containing proteins. Homo- and hetero-oligomerisation could play a part in regulating the localization of Zfp647 and other KRAB zinc finger proteins, therefore identifying the domain and amino acid residues responsible for interaction could prove to be important for other future projects. Likewise, it would be interesting to find out whether Jmjd1c and Jarid1b can bind to KRAB zinc finger proteins directly, as witnessed with Zkscan17 and JARID2 (Mysliwiec *et al*, 2007), and also whether these proteins can bind to other KRAB zinc finger protein repressors such as KAP-1 and HP1. If the jumonji proteins can bind to KAP-1, this may point to these proteins being general KRAB-mediated transcriptional repressors, rather than effector proteins recruited by particular KRAB zinc finger proteins to a few specific genes. As I have seen that NEM treatment may affect Jarid1b's interaction with Zfp647, it would also be of worth to couple the investigations into Zfp647's modification and protein-protein interactions, to see whether the prevention of Zfp647 modification by amino acid mutation also enhances or prevents its interaction with any potential interactors.

The nature of the modification of Zfp647 has not been clearly identified in this thesis. Evidence suggests that the protein can be ubiquitylated, however this remains to be verified. The next step to be taken in proving that this is the case, as outlined earlier, is to test Zfp647 in an *in vitro* ubiquitylation assay. If Zfp647 does not appear to be ubiquitylated by this assay, it could also be tested for sumoylation or

neddylation in a similar manner, or by mass spectrometry. Once the species and location of modification are identified, the effect of modification on Zfp647's function can be tested. KAP-1 sumoylation is known to be enhanced by interaction with KRAB zinc finger proteins (Mascle *et al*, 2007), and is required for its interaction with Mi-2_ and SETDB1 (Ivanov *et al*, 2007), therefore it is possible that modification of KRAB zinc finger proteins regulates their interaction with KAP-1 or other interacting proteins. If Zfp647 modification enhances or is required for interaction with KAP-1, this may explain why the signal achieved from my *in vitro* co-immunoprecipitation experiment was weak. Given that we have previously demonstrated that over-expressed Zfp647 mis-localises in tissue culture cells, it may be difficult to study the role Zfp647 modification plays in its localisation. However, if a Zfp647 construct could be generated, possibly one that is expressed at a low level, that shows the same localisation as the endogenous protein in pMEF cells, mutation of amino acid residues that are critical for modification could be used to analyse the effect of loss of modification on the formation of the KAKA foci, to see whether modification is important for the recruitment of KRAB zinc finger proteins to these structures.

Following these experiments to verify the results I have presented in this thesis, one of the most important steps would be to identify the gene targets of Zfp647. Currently, the materials that we have available are a gene-trapped mouse line, and two antibodies that can both bind to endogenous Zfp647. Gene targets could be identified in a number of ways: expression analysis of gene-trapped homozygous mice versus their wild-type littermates by microarray (Lockhart *et al*, 1996); ChIP-on-chip experiments using either or both of the antibodies on whole genome arrays (Weinmann *et al*, 2002); or ChIP-seq experiments using one of the new generation of whole-genome sequencing machines (Johnson *et al*, 2007). The discovery of Zfp647's gene targets would make it one of very few KRAB zinc finger proteins whose binding sites in the genome were known, and would allow the further investigation of the KAKA foci. For example, it would allow the study of the localization of the gene targets in relation to the KAKA foci in undifferentiated and differentiated ES cells. If these sequences were found to be recruited alongside

Zfp647 to the KAKA foci, it would suggest that KAKA foci could be functioning as a silent transcriptional environment. This could be backed up by qRT-PCR analysis of the expression of the targets in cells at both differentiation states. Most studies looking at the repression (or activation) of transcription stimulated by KRAB zinc finger proteins have generally fused either the KRAB domain or full-length protein to a DNA-binding domain of another transcription factor, thus tethering it to a reporter gene. Simple ChIP-PCR experiments alongside expression analysis in cells where gene targets are either bound by or free from Zfp647 would therefore be one of the first demonstrations of the effect of binding of a KRAB zinc-finger protein to its endogenous gene targets.

As discussed in Section 1.3.6, the KRAB zinc finger protein genes have been shown to be bound *in vivo* by both KAP-1 and HP1β (O'Geen *et al*, 2007; Vogel *et al*, 2006). KAP-1 binding sites were typically ~800bp long, whereas the domains bound by HP1 were much larger, at up to 4Mb in length. Vogel *et al* did not find significant downregulation in the expression of KRAB zinc finger protein genes that were bound by HP1β, and therefore speculated that the heterochromatinisation of these genes was perhaps a method to prevent their HR-driven excision from the genome (Vogel *et al*, 2006). However, O'Geen *et al* reported that those proteins subjected to KAP-1 binding and histone H3K9 methylation were expressed at lower levels compared to all other expressed genes, and stated that they had found that particular KRAB zinc finger proteins were expressed at high levels in the Ntera2 cells used in their experiments. Therefore they predicted that the KRAB zinc finger protein genes are auto-regulated by certain KRAB zinc finger proteins and KAP-1. It is therefore possible that the KAKA foci that we see in OS25 cells are domains of KRAB zinc finger protein genes bound by certain KRAB zinc finger proteins, KAP-1 and HP1. Fluorescence *in situ* hybridization could be carried out on the cells using the three KRAB zinc finger protein genes (*ZNF333*, *ZNF426* and *ZNF554*) described by O'Geen *et al* that are silent in Ntera2 cells to show their localization, and this technique could be carried out alongside immunofluorescence for KAP-1 and HP1 to see whether the genes co-localise with these proteins at the non-heterochromatic

KAKA foci. If the master-regulator KRAB zinc finger proteins implied by O'Geen *et al* were also known, then their proximity to these genes could also be assayed.

Finally, we have found that KAKA foci are situated adjacent to PML-NBs in ES cells (Briers *et al*, 2009) and in pMEF cells. The role of PML in this localization could be investigated by the study of Zfp647 in pMEFs lacking PML (Wang *et al*, 1998). If the KAKA foci failed to form in these cells, or their localization was different in another way to that in wild-type cells, this would suggest that the PML network is required for the anchoring of the KAKA foci. Additionally, if it is discovered that Zfp647 is sumoylated, the localization of the KAKA foci could also be looked at in cells where PML's SIM is mutated, as this also has a profound effect on the structure of PML-NBs, and could in turn influence KAKA foci formation (Lin *et al*, 2006; Shen *et al*, 2006). The effect of loss of PML on Zfp647 gene target expression and Zfp647 post-translational modification could also be analysed in these cells.

References

Abrink, M., Ortiz, J.A., Mark, C., Sanchez, C., Looman, C., Hellman, L., Chambon, P. and Losson, R. (2001) Conserved interaction between distinct Krüppel-associated box domains and the transcriptional intermediary factor 1 _. *Proceedings of the National Academy of Sciences USA*, 98: 1422-6.

Agata, Y., Matsuda, E. and Shimizu, A. (1999) Two novel Krüppel-associated box-containing zinc-finger proteins, KRAZ1 and KRAZ2, repress transcription through functional interaction with the corepressor KAP-1 (TIF1_/KRIP-1). *Journal of Biological Chemistry*, 274: 16412-22.

Agger, K., Cloos, P.A., Christensen, J., Pasini, D., Rose, S., Rappsilber, J., Issaeva, I., Canaani, E., Salcini, A.E. and Helin, K. (2007) UTX and JMJD3 are histone H3K27 demethylases involved in HOX gene regulation and development. *Nature*, 449: 731-4.

Ascoli, C.A. and Maul, G.G. (1991) Identification of a novel nuclear domain. *Journal of Cell Biology*, 112: 785-95.

Asmuß, M., Mullenders, L.H.F. and Hartwig, A. (2000) Interference by toxic metal compounds with isolated zinc finger DNA repair proteins. *Toxicology Letters*, 112-113: 227-31.

Ausió, J., Abbott, D.W., Wang, X. and Moore, S.C. (2001) Histone variants and histone modifications: a structural perspective. *Biochemistry and Cell Biology*, 79:693-708.

Ayyanathan, K., Lechner, M.S., Bell, P., Maul, G.G., Schultz, D.C., Yamada, Y., Tanaka, K., Torigoe, K. and Rauscher, F.J.3rd (2003) Regulated recruitment of HP1 to a euchromatic gene induces mitotically heritable, epigenetic gene silencing: a mammalian cell culture model of gene variegation. *Genes and Development*, 17: 1855-69.

Bannister, A.J., Zegerman, P., Partridge, J.F., Miska, E.A., Thomas, J.O., Allshire, R.C. and Kouzarides, T. (2001) Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature*, 410: 120-4.

Barski, A., Cuddapah, S., Cui, K., Roh, T.Y., Schones, D.E., Wang, Z., Wei, G., Chepelev, I. and Zhao, K. (2007) High-resolution profiling of histone methylations in the human genome. *Cell*, 129: 823-37.

Bates, D.L. and Thomas, J.O. (1981) Histones H1 and H5: one or two molecules per nucleosome? *Nucleic Acids Research*, 9: 5883-94.

Bellefroid, E.J., Poncelet, D.A., Lecocq, P.J., Revelant, O. and Martial, J.A. (1991) The evolutionarily conserved Krüppel-associated box domain defines a subfamily of eukaryotic multifingered proteins. *Proceedings of the National Academy of Sciences USA*, 88: 3608-12.

Bellefroid, E.J. *et al* (1993) Clustered organization of homologous KRAB zinc-finger genes with enhanced expression in human T lymphoid cells. *EMBO Journal*, 12: 1363-74.

Bender, M.A., Bulger, M., Close, J. and Groudine, M. (2000) _-globin gene switching and DNase I sensitivity of the endogenous _-globin locus in mice do not require the locus control region. *Molecular Cell*, 5: 387-93.

Bernstein, B.E. *et al* (2005) Genomic maps and comparative analysis of histone modifications in human and mouse. *Cell*, 120: 169-81.

Bernstein, B.E. *et al* (2006) A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell*, 125: 315-26.

Best, J.L., Ganiatsas, S., Agarwal, S., Changou, A., Salomoni, P., Shirihai, O., Meluh, P.B., Pandolfi, P.P. and Zon, L.I. (2002) SUMO-1 protease-1 regulates gene transcription through PML. *Molecular Cell*, 10: 843-55.

Billon, N., Jolicouer, C., Ying, Q.L., Smith, A. and Raff, M. (2002) Normal timing of oligodendrocyte development from genetically engineered, lineage-selectable mouse ES cells. *Journal of Cell Science,* 115: 3657-65.

Bird, A., Taggart, M., Frommer, M., Miller, O.J. and Macleod, D. (1985) A fraction of the mouse genome that is derived from islands of nonmethylated, CpG-rich DNA. *Cell*, 40: 91-9.

Bischoff, K.R., Klebe, C., Kretschmer, J., Wittinghofer, A. and Ponstingl, H. (1994) RanGAP1 induces GTPase activity of nuclear Ras-related Ran. *Proceedings of the National Academy of Sciences USA*, 91: 2587-91.

Blencowe, B.J., Nickerson, J.A., Issner, R., Penman, S. and Sharp, P.A. (1994) Association of nuclear matrix antigens with exon-containing splicing complexes. *Journal of Cell Biology*, 127: 593-607.

Boddy, M.N., Howe, K., Etkin, L.D., Solomon, E. and Freemont, P.S. (1996) PIC1, a novel ubiquitin-like protein which interacts with the PML component of a multiprotein complex that is disrupted in acute promyelocytic leukaemia. *Oncogene*, 13: 971-82.

Boddy, M.N., Duprez, E., Borden, K.L. and Freemont, P.S. (1997) Surface residue mutations of the PML RING finger domain alter the formation of nuclear matrix-associated PML bodies. *Journal of Cell Science*, 110: 2197-205.

Bøe, S.O., Haave, M., Jul-Larsen, A., Grudic, A., Bjerkvig, R. and Lønning, P.E. (2006) Promyelocytic leukemia nuclear bodies are predetermined processing sites for damaged DNA. *Journal of Cell Science*, 119: 3284-95.

Boisvert, F.M., Hendzel, M.J. and Bazett-Jones, D.P. (2000) Promyelocytic leukemia (PML) nuclear bodies are protein structures that do not accumulate RNA. *Journal of Cell Biology*, 148: 283-92.

Bolzer, A., Kreth, G., Solovei, I., Koehler, D., Saracoglu, K., Fauth, C., Müller, S., Eils, R., Cremer, C., Speicher, M.R. and Cremer, T. (2005) Three-dimensional maps of all chromosomes in human male fibroblast nuclei and prometaphase rosettes. *PLoS Biology*, 3: e157.

Borden, K.L., Lally, J.M., Martin, S.R., O'Reilly, N.J., Etkin, L.D. and Freemont, P.S. (1995) Novel topology of a zinc-binding domain from a protein involved in regulating early *Xenopus* development. *EMBO Journal*, 14: 5947-56.

Borden, K.L., Lally, J.M., Martin, S.R., O'Reilly, N.J., Solomon, E. and Freemont, P.S. (1996) *In vivo* and *in vitro* characterization of the B1 and B2 zinc-binding domains from the acute promyelocytic leukemia protooncoprotein PML. *Proceedings of the National Academy of Sciences USA*, 93: 1601-6.

Boyer, L.A. *et al* (2006) Polycomb complexes repress developmental regulators in murine embryonic stem cells. *Nature*, 441: 349-53.

Boyle, S., Gilchrist, S., Bridger, J.M., Mahy, N.L., Ellis, J.A. and Bickmore, W.A. (2001) The spatial organization of human chromosomes within the nuclei of normal and emerin-mutant cells. *Human Molecular Genetics*, 10: 211-9.

Branco, M.R. and Pombo, A. (2006) Intermingling of chromosome territories in interphase suggests role in translocations and transcription-dependent associations. *PLoS Biology*, 4: e138.

Briers, S. (2005) Characterisation of KRAB-ZFPs in the mouse. *Thesis presented for the Degree of Doctor of Philosophy, University of Edinburgh*.

Briers, S., Crawford, C., Bickmore, W.A. and Sutherland, H.G. (2009) KRAB zinc-finger proteins localise to novel KAP1-containing foci that are adjacent to PML nuclear bodies. *Journal of Cell Science*, 122: 937-46.

Brown, K.E., Guest, S.S., Smale, S.T., Hahm, K., Merkenschlager, M. and Fisher, A.G. (1997) Association of transcriptionally silent genes with Ikaros complexes at centromeric heterochromatin. *Cell*, 91: 845-54.

Brown, K.E., Baxter, J., Graf, D., Merkenschlager, M. and Fisher, A.G. (1999) Dynamic repositioning of genes in the nucleus of lymphocytes preparing for cell division. *Molecular Cell*, 3: 207-17.

Brown, K.E., Amoils, S., Horn, J.M., Buckle, V.J., Higgs, D.R., Merkenschlager, M. and Fisher, A.G. (2001) Expression of _- and _-globin genes occurs within different nuclear domains in haemopoietic cells. *Nature Cell Biology*, 3: 602-6.

Brown, J.M., Leach, J., Reittie, J.E., Atzberger, A., Lee-Prudhoe, J., Wood, W.G., Higgs, D.R., Iborra, F.J. and Buckle, V.J. (2006) Coregulated hman globin genes are frequently in spatial proximity when active. *Journal of Cell Biology*, 172: 177-87.

Brown, J.M., Green, J., das Neves, R.P., Wallace, H.A., Smith, A.J., Hughes, J., Gray, N., Taylor, S., Wood, W.G., Higgs, D.R., Iborra, F.J. and Buckle, V.J. (2008) Association between active genes occurs at nuclear speckles and is modulated by chromatin environment. *Journal of Cell Biology*, 182: 1083-97.

Cáceres, J.F., Misteli, T., Screaton, G.R., Spector, D.L. and Krainer, A.R. (1997) Role of the modular domains of SR proteins in subnuclear localization and alternative splicing specificity. *Journal of Cell Biology*, 138: 225-38.

Cammas, F., Mark, M., Dollé, P., Dierich, A., Chambon, P. and Losson, R. (2000) Mice lacking the transcriptional corepressor TIF1_ are defective in early postimplantation development. *Development*, 127: 2955-63.

Cammas, F., Oulad-Abdelghani, M., Vonesch, J.L., Huss-Garcia, Y., Chambon, P. and Losson, R. (2002) Cell differentiation induces TIF1_ association with centromeric heterochromatin via an HP1 interaction. *Journal of Cell Science*, 115: 3439-48.

Cammas, F., Herzog, M., Lerouge, T., Chambon, P. and Losson, R. (2004) Association of the transcriptional corepressor TIF1_ with heterochromatin protein 1 (HP1): an essential role for progression through differentiation. *Genes and Development*, 18: 2147-60.

Camporeale, G., Shubert, E.E., Sarath, G., Cerny, R. and Zempleni, J. (2004) K8 and K12 are biotinylated in human histone H4. *European Journal of Biochemistry*, 271: 2257-63.

Cao, R., Wang, L., Wang, H., Xia, L., Erdjument-Bromage, H., Tempst, P., Jones, R.S. and Zhang, Y. (2002) Role of histone H3 lysine 27 methylation in Polycomb-group silencing. *Science*, 298: 1039-43.

Cao, L., Wang, Z., Zhu, C., Zhao, Y., Yuan, W., Li, J., Wang, Y., Ying, Z., Li, Y., Yu, W., Wu, X. and Liu, M. (2005) ZNF383, a novel KRAB-containing zinc finger protein, suppresses MAPK signalling pathway. *Biochemical and Biophysical Research Communications*, 333: 1050-9.

Carter, K.C., Taneja, K.L. and Lawrence, J.B. (1991) Discrete nuclear domains of poly(A) RNA and their relationship to the functional organization of the nucleus. *Journal of Cell Biology*, 115: 1191-202.

Casademunt, E., Carter, B.D., Benzel, I., Frade, J.M., Dechant, G. and Barde, Y.A. (1999) The zinc finger protein NRIF interacts with the neurotrophin receptor $p75^{NTR}$ and participates in programmed cell death. *EMBO Journal*, 18: 6050-61.

Chalkiadaki, A. and Talianidis, I. (2005) SUMO-dependent compartmentalization in promyelocytic leukemia protein nuclear bodies prevents the access of LRH-1 to chromatin. *Molecular and Cellular Biology*, 25: 5095-105.

Chambeyron, S. and Bickmore, W.A. (2004) Chromatin decondensation and nuclear reorganization of the HoxB locus upon induction of transcription. *Genes and Development*, 18: 1119-30.

Chang, B., Chen, Y., Zhao, Y. and Bruick, R.K. (2007) JMJD6 is a histone arginine demethylase. *Science*, 318: 444-7.

Chang, C.W., Chou, H.Y., Lin, Y.S., Huang, K.H., Chang, C.J., Hsu, T.C. and Lee, S.C. (2008) Phosphorylation at Ser473 regulates heterochromatin protein 1 binding and corepressor function of TIF1_/KAP1. *BMC Molecular Biology*, 9: 61.

Chen, D., Zhang, Z., Li, M., Wang, W., Li, Y., Rayburn, E.R., Hill, D.L., Wang, H. and Zhang, R. (2007) Ribosomal protein S& as a novel modulator of p53-MDM2 interaction: binding to MDM2, stabilization of p53 protein, and activation of p53 function. *Oncogene*, 26: 5029-37.

Christensen, J., Agger, K., Cloos, P.A., Pasini, D., Rose, S., Sennels, L., Rappsilber, J., Hansen, K.H., Salcini, A.E. and Helin, K. (2007) RBP2 belongs to a family of demethylases, specific for tri- and dimethylated lysine 4 on histone 3. *Cell*, 128: 1063-76.

Chung, H.R., Schäfer, U., Jäckle, H. and Böhm, S. (2002) Genomic expansion and clustering of ZAD-containing C2H2 zinc-finger genes in *Drosophila*. *EMBO Reports*, 3: 1158-62.

Citterio, E., Papait, R., Nicassio, F., Vecchi, M., Gomiero, P., Mantovani, R., Di Fiore, P.P. and Bonapace, I.M. (2004) Np95 is a histone-binding protein endowed with ubiquitin ligase activity. *Molecular and Cellular Biology*, 24: 2526-35.

Clayton, A.L., Rose, S., Barratt, M.J. and Mahadevan, L.C. (2000) Phosphoacetylation of hostone H3 on c-fos and c-jun-associated nucleosomes upon gene activation. *EMBO Journal*, 19: 3714-26.

Clayton, A.L. and Mahadevan, L.C. (2003) MAP kinase-mediated phosphoacetylation of histone H3 and inducible gene regulation. *FEBS Letters*, 546: 51-8.

Cloos, P.A., Christensen, J., Agger, K., Maiolica, A., Rappsilber, J., Antal, T., Hansen, K.H. and Helin, K. (2006) The putative oncogene GASC1 demethylates tri- and dimethylated lysine 9 on histone H3. *Nature*, 442: 307-11.

Cloos, P.A., Christensen, J., Agger, K. and Helin, K. (2008) Erasing the methyl mark: histone demethylases at the center of cellular differentiation and disease. *Genes and Development*, 22: 1115-40.

Cobb, B.S., Morales-Alcelay, S., Kleiger, G., Brown, K.E., Fisher, A.G. and Smale, S.T. (2000) Targeting of Ikaros to pericentromeric heterochromatin by direct DNA binding. *Genes and Development*, 14: 2146-60.

Cohen-Fix, O., Peters, J.M., Kirschner, M.W. and Koshland, D. (1996) Anaphase initiation in *Saccharomyces cerevisiae* is controlled by the APC-dependent degradation of the anaphase inhibitor Pds1p. *Genes and Development*, 10: 3081-93.

Comings, D.E. (1980) Arrangement of chromatin in the nucleus. *Human Genetics*, 53: 131-43.

Contreras, A., Hale, T.K., Stenoien, D.L., Rosen, J.M., Mancini, M.A. and Herrera, R.E. (2003) The dynamic mobility of histone H1 is regulated by cyclin/CDK phosphorylation. *Molecular and Cellular Biology*, 23: 8626-36.

Cowieson, N.P., Partridge, J.F., Allshire, R.C. and McLaughlin, P.J. (2000) Dimerisation of a chromo shadow domain and distinctions from the chromodomain as revealed by structural analysis. *Current Biology*, 10: 517-25.

Craig, J.M. and Bickmore, W.A. (1994) The distribution of CpG islands in mammalian chromosomes. *Nature Genetics*, 7: 376-82.

Croft, J.A., Bridger, J.M., Boyle, S., Perry, P., Teague, P. and Bickmore, W.A. (1999) Differences in the localization and morphology of chromosomes in the human nucleus. *Journal of Cell Biology*, 145: 1119-31.

Cunliffe, V.T. (2008) Eloquent silence: developmental functions of Class I histone deacetylases. *Current Opinion in Genetics and Development*, 18: 404-10.

Czermin, B., Melfi, R., McCabe, D., Seitz, V., Imhof, A. and Pirrotta, V. (2002) Drosophila enhancer of Zeste/ESC complexes have a histone H3 methyltransferase activity that marks chromosomal Polycomb sites. *Cell*, 111: 185-96.

D'Orazi, G. *et al* (2002) Homeodomain-interacting protein kinase-2 phosphorylates p53 at Ser 46 and mediates apoptosis. *Nature Cell Biology*, 4: 11-9.

Davidovic, L., Vodenicharov, M., Affar, E.B. and Poirier, G.G. (2001) Importance of poly(ADP-ribose) glycohydrolase in the control of poly(ADP-ribose) metabolism. *Experimental Cell Research*, 268: 7-13.

De Napoles, M., Mermoud, J.E., Wakao, R., Tang, Y.A., Endoh, M., Appanah, R., Nesterova, T.B., Silva, J., Otte, A.P., Vidal, M., Koseki, H. and Brockdorff, N. (2004) Polycomb group proteins Ring1A/B link ubiquitylation of histone H2A to heritable gene silencing and X inactivation. *Developmental Cell*, 7: 663-76.

De Santa, F., Totaro, M.G., Prosperini, E., Notarbartolo, S., Testa, G. and Natoli, G. (2007) The histone H3 lysine-27 demethylase Jmjd3 links inflammation to inhibition of polycomb-mediated gene silencing. *Cell*, 130: 1083-94.

De Thé, H., Lavau, C., Marchio, A., Chomienne, C., Degos, L. and Dejean, A. (1991) The PML-RAR_ fusion mRNA generated by the t(15:17) translocation in acute promyelocytic leukaemia encodes a functionally altered RAR. *Cell*, 66: 675-84.

Dellaire, G. and Bazett-Jones, D.P. (2004) PML nuclear bodies: dynamic sensors of DNA damage and cellular stress. *Bioessays*, 26: 963-77.

Dellaire, G., Ching, R.W., Ahmed, K., Jalali, F., Tse, K.C., Bristow, R.G. and Bazett-Jones, D.P. (2006) Promyelocytic leukemia nuclear bodies behave as DNA damage sensors whose response to DNA double-strand breaks is regulated by NBS1 and the kinases ATM and ATR. *Journal of Cell Biology*, 175: 55-66.

Denslow, S.A. and Wade, P.A. (2007) The human Mi-2/NuRD complex and gene regulation. *Oncogene*, 26: 5433-8.

Deuring, R. *et al* (2000) The ISWI chromatin-remodelling protein is required for gene expression and the maintenance of higher order chromatin structure *in vivo*. *Molecular Cell*, 5: 355-65.

Denslow, S.A. and Wade, P.A. (2007) The human Mi-2/NuRD complex and gene regulation. *Oncogene*, 26: 5433-8.

Devault, A., Martinez, A.M., Fesquet, D., Labbé, J.C., Morin, N., Tassan, J.P., Nigg, E.A., Cavadore, J.C. and Doreé, M. (1995) MAT1 ('menage à trois') a new RING finger protein subunit stabilizing cyclin H-cdk7 complexes in starfish and Xenopus CAK. *EMBO Journal*, 14: 5027-36.

Dhalluin, C., Carlson, J.E., Zeng, L., He, C., Aggarwal, A.K. and Zhou, M.M. (1999) Structure and ligand of a histone acetyltransferase bromodomain. *Nature*, 399: 491-6.

Ding, W., Liu, W., Cooper, K.L., Qin, X., Bergo, P.L., Hudson, L.G. and Liu, K.J. (2008) Inhibition of PARP-1 by arsenite interferes with repair of oxidative DNA damage. *Journal of Biological Chemistry*, Epub Manuscript M805566200.

Dorigo, B., Schalch, T., Kulangara, A., Duda, S., Schroeder, R.R. and Richmond, T.J. (2004) Nucleosome arrays reveal the two-start organization of the chromatin fiber. *Science*, 306: 1571-3.

Dillon, S.C., Zhang, X., Trievel, R.C. and Cheng, X. (2005) The SET-domain protein superfamily: protein lysine methyltransferases. *Genome Biology*, 6: 227

Drissen, R., Palstra, R.J., Gillemans, N., Splinter, E., Grosveld, F., Philipsen, S. and de Laat, W. (2004) The active spatial organization of the _-globin locus requires the transcription factor EKLF. *Genes and Development*, 18: 2485-90.

Duan, Z., Person, R.E., Lee, H.H., Huang, S., Donadieu, J., Badolato, R., Grimes, H.L., Papayannopoulou, T. And Horwitz, M.S. (2007) Epigenetic regulation of protein-coding and microRNA genes by the Gfi1-interacting tumor suppressor PRDM5. *Molecular and Cellular Biology*, 27: 6889-902.

Dyck, J.A., Maul, G.G., Miller, W.H.Jr, Chen, J.D., Kakizuka, A. and Evans, R.M. (1994) A novel macromolecular structure is a target of the promyelocyte-retinoic acid receptor oncoprotein. *Cell*, 76: 333-43.

Dyson, M.H., Thomson, S., Inagaki, M., Goto, H., Arthur, S.J., Nightingale, K., Iborra, F.J. and Mahadevan, L.C. (2005) MAP kinase-mediated phosphorylation of distinct pools of histone H3 at S10 or S28 via mitogen- and stress-activated kinase 1/2. *Journal of Cell Science*, 118: 2247-59.

Eisen, J.A., Sweder, K.S. and Hanawalt, P.C. (1995) Evolution of the SNF2 family of proteins: subfamilies with distinct sequences and functions. *Nucleic Acids Research*, 23: 2715-23.

Eskiw, C.H., Dellaire, G., Mymryk, J.S. and Bazett-Jones, D.P. (2003) Size, position and dynamic behaviour of PML nuclear bodies following cell stress as a paradigm for supramolecular trafficking and assembly. *Journal of Cell Science*, 116: 4455-66.

Evdokimov, E., Sharma, P., Lockett, S.J., Lualdi, M. and Kuehn, M.R. (2008) Loss of SUMO1 in mice affects RanGap1 localization and formation of PML nuclear bodies, but is not lethal as it can be compensated by SUMO2 or SUMO3. *Journal of Cell Science*, 121: 4106-4113.

Everett, R.D., Earnshaw, W.C., Pluta, A.F., Sternsdorf, T., Ainsztein, A.M., Carmena, M., Ruchaud, S., Hsu, W.L. and Orr, A. (1999) A dynamic connection between centromeres and ND10 proteins. *Journal of Cell Science*, 112: 3443-54.

Fahmy, O.G. and Fahmy, M. (1959) New mutants report. *Drosophila Information Service*, 33: 82-94.

Fan, Y., Nikitina, T., Zhao, J., Fleury, T.J., Bhattacharyya, R., Bouhassira, E.E., Stein, A., Woodcock, C.L. and Skoultchi, A.I. (2005) Histone H1 depletion in mammals alters global chromatin structure but causes specific changes in gene regulation. *Cell*, 123: 1199-212.

Felsenfeld, G. and Groudine, M. (2003) Controlling the double helix. *Nature*, 421: 448-53.

Feng, Q. and Zhang, Y. (2001) The MeCP1 complex represses transcription through preferential binding, remodelling, and deacetylating methylated nucleosomes. *Genes and Development*, 15: 827-32.

Filion, G.J., Zhenilo, S., Salozhin, S., Yamada, D., Prokhortchouk, E. and Defossez, P.A. (2006) A family of human zinc finger proteins that bind methylated DNA and repress transcription. *Molecular and Cellular Biology*, 26: 169-81.

Finlan, L.E., Sproul, D., Thomson, I., Boyle, S., Kerr, E., Perry, P., Ylstra, B., Chubb, J.R. and Bickmore, W.A. (2008) Recruitment to the nuclear periphery can alter expression of genes in human cells. *PLoS Genetics*, 4: e1000039.

Fischle, W., Wang, Y., Jacobs, S.A., Kim, Y., Allis, C.D. and Khorasanizadeh, S. (2003) Molecular basis for the discrimination of repressive methyl-lysine marks in histone H3 by Polycomb and HP1 chromodomains. *Genes and Development*, 17: 1870-81.

Fisher, R.P., Jin, P., Chamberlin, H.M. and Morgan, D.O. (1995) Alternative mechanisms of CAK assembly require an assembly factor or an activating kinase. *Cell*, 83: 47-57.

Fleischer, S.,Wiemann, S., Hill, H. and Hofmann, T.G. (2006) PML-associated repressor of transcription (PAROT), a novel KRAB-zinc finger repressor, is regulated through association with PML nuclear bodies. *Experimental Cell Research*, 312: 901-12.

Fodor, B.D., Kubichek, S., Yonezawa, M., O'Sullivan, R.J., Sengupta, R., Perez-Burgos, L., Opravil, S., Mechtler, K., Schotta, G. and Jenuwein, T. (2006) Jmjd2b antagonizes H3K9 trimethylation at pericentric heterochromatin in mammalian cells. *Genes and Development*, 20: 1557-62.

Francis, N.J., Kingston, R.E. and Woodcock, C.L. (2004) Chromatin compaction by a polycomb protein complex. *Science*, 306: 1574-7.

Friedman, J.R., Fredericks, W.J., Jensen, D.E., Speicher, D.W., Huang, X.P., Neilson, E.G. and Rauscher, F.J.3rd (1996) KAP-1, a novel corepressor for the highly conserved KRAB repression domain. *Genes and Development*, 10: 2067-78.

Fu, X.D. and Maniatis, T. (1990) Factor required for mammalian spliceosome assembly is localized to discrete regions in the nucleus. *Nature*, 343: 437-41.

Gallo, A., Keegan, L.P., Ring, G.M. and O'Connell, M.A. (2003) An ADAR that edits transcripts encoding ion channel subunits functions as a dimer. *EMBO Journal*, 22: 3421-30.

Gentry, J.J., Rutkoski, N.J., Burke, T.L. and Carter, B.D. (2004) A functional interaction between the p75 neurotrophin receptor interacting factors, TRAF6 and NRIF. *Journal of Biological Chemistry*, 279: 16646-56.

Georgatsou, E., Bourgarel, P. and Meo, T. (1993) Male-specific expression of mouse sex-limited protein requires growth hormone, not testosterone. *Proceedings of the National Academy of Sciences USA*, 90: 3626-30.

Germain-Desprez, D., Bazinet, M., Bouvier, M. and Aubry, M. (2003) Oligomerization of transcriptional intermediary factor 1 regulators and interaction with ZNF74 nuclear matrix protein revealed by bioluminescence resonance energy transfer in living cells. *Journal of Biological Chemistry*, 278: 22367-73.

Geulen, L., Pagie, L., Brasset, E., Meuleman, W., Faza, M.B., Talhout, W., Eussen, B.H., de Klein, A., Wessels, L., de Laat, W. and van Steensel, B. (2008) Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. *Nature*, 453: 948-51.

Gilbert, N., Boyle, S., Fiegler, H., Woodfine, K., Carter, N.P. and Bickmore, W.A. (2004) Chromatin architecture of the human genome; gene-rich domains are enriched in open chromatin fibers. *Cell*, 118: 555-66.

Girdwood, D.W., Tatham, M.H. and Hay, R.T. (2004) SUMO and transcriptional regulation. *Seminars in Cell and Developmental Biology*, 15: 201-10.

Goldmark, J.P., Fazzio, T.G., Estep, P.W., Church, G.M. and Tsukiyama, T. (2000) The Isw2 chromatin remodelling complex represses early meiotic genes upon recruitment by Ume6p. *Cell*, 103: 423-33.

Hamel, P.A., Gill, R.M., Phillips, R.A. and Gaillie, B.L. (1992) Transcriptional repression of the E2-containing promoters EllaE, c-myc, and RB1 by the product of the RB1 gene. *Molecular and Cellular Biology*, 12: 3431-8.

Hamilton, A.T., Huntley, S., Tran-Gyamfi, M., Baggott, D.M., Gordon, L. and Stubbs, L. (2006) Evolutionary expansion and divergence in the ZNF91 subfamily of primate-specific zinc finger genes. *Genome Research*, 16: 584-94.

Hang, J. and Dasso, M. (2002) Association of the human SUMO-1 protease SENP2 with the nuclear pore. *Journal of Biological Chemistry*, 277: 19961-6.

Harlow, E. and Lane, D. (1988) Antibodies: a laboratory manual. *Cold Spring Harbor Laboratories Press*.

Hartwig, A. (1998) Carcinogenicity of metal compounds: possible role of DNA repair inhibition. *Toxicology Letters*, 102-103: 235-9.

Hassa, P.O., Buerki, C., Lombardi, C., Imhof, R. and Hottiger, M.O. (2003) Transcriptional coactivation of nuclear factor-_B-dependent gene expression by p300 is regulated by poly(ADP)-ribose polymerase-1. *Journal of Biological Chemistry*, 278: 45145-53.

Hassa, P.O. and Hottiger, M.O. (2008) The diverse biological roles of mammalian PARPs, a small but powerful family of poly-ADP-ribose polymerases. *Frontiers in Bioscience*, 13: 3046-82.

Hay, R.T. (2005) SUMO: a history of modification. *Molecular Cell*, 18: 1-12.

Hazzalin, C.A. and Mahadevan, L.C. (2005) Dynamic acetylation of all lysine 4-methylated histone H3 in the mouse nucleus: analysis at c-fos and c-jun. *PLoS Biology*, 3: e393.

Hebbes, T.R., Thorne, A.W. and Crane-Robinson, C. (1988) A direct link between core histone acetylation and transcriptionally active chromatin. *EMBO Journal*, 7: 1395-402.

Hendrich, B. and Bird, A. (1998) Identification and characterization of a family of mammalian methyl-CpG binding proteins. *Molecular and Cellular Biology*, 18: 6538-47.

Henneman, H., Vassen, L., Geisen, C., Eilers, M. and Möröy, T. (2003) Identification of a novel Krüppel-associated box domain protein, KRIM-1, that interacts with c-myc and inhibits its oncogenic activity. *Journal of Biological Chemistry*, 278: 28799-811.

Hershko, A. and Ciechanover, A. (1998) The ubiquitin system. *Annual Reviews of Biochemistry*, 67: 425-79.

Hewitt, S.L., High, F.A., Reiner, S.L., Fisher, A.G. and Merkenschlager, M. (2004) Nuclear repositioning marks the selective exclusion of lineage-inappropriate transcription factor loci during T helper cell differentiation. *European Journal of Immunology*, 34: 3604-13.

Hiebert, S.W., Chellappan, S.P., Horowitz, J.M. and Nevins, J.R. (1992) The interaction of RB with E2F coincides with an inhibition of the transcriptional activity of E2F. *Genes and Development*, 6: 177-85.

Hofman, K., Swinnen, J.V., Claessens, F., Verhoeven, G. and Heyns, W. (2003) The retinoblastoma protein-associated transcription repressor RBaK interacts with the androgen receptor and enhances its transcriptional activity. *Journal of Molecular Endocrinology*, 31: 583-96.

Holaska, J.M. and Wilson, K.L. (2007) An emerin "proteome": purification of distinct emerin-containing complexes from HeLa cells suggests molecular basis for diverse roles including gene regulation, mRNA splicing, signalling, mechanosensing, and nuclear architecture. *Biochemistry*, 46: 8897-908.

Honer, C., Chen, P., Toth, M.J. and Schumacher, C. (2001) Identification of SCAN dimerization in four gene families. *Biochimica et Biophysica Acta*, 1517: 441-8.

Hong, L., Schroth, G.P., Matthews, H.R., Yau, P. And Bradbury, E.M. (1993) Studies of the DNA binding properties of histone H4 amino terminus. Thermal denaturation studies reveal that acetylation markedly reduces the binding constant of the H4 "tail" to DNA. *Journal of Biological Chemistry*, 268: 305-14.

Horsley, D., Hutchings, A., Butcher, G.W. and Singh, P.B. (1996) M32, a murine homologue of *Drosophila* heterochromatin protein 1 (HP1), localises to euchromatin within interphase nuclei and is largely excluded from constitutive heterochromatin. *Cytogenetics and Cell Genetics*, 73: 308-11.

Hoshino, S., Kikkawa, S., Takahashi, K., Itoh, H., Kaziro, Y., Kawasaki, H., Suzuki, K., Katada, T. and Ui, M. (1990) Identification of sites for alkylation by N-ethylmaleimide and pertussis toxin-catalyzed ADP-ribosylation on GTP-binding proteins. *FEBS Letters*, 276: 227-31.

Huang, S., Litt, M., Felsenfeld, G. (2005) Methylation of histone H4 by arginine methyltransferase PRMT1 is essential *in vivo* for many subsequent histone modifications. *Genes and Development*, 19: 1885-93.

Huang, Y., Fang, J., Bedford, M.T., Zhang, Y. and Xu, R.M. (2006) Recognition of histone H3 lysine-4 methylation by the double tudor domain of JMJD2A. *Science*, 312: 748-51.

Huang, C., Jia, Y., Yang, S., Chen, B., Sun, H., Shen, F. and Wang, Y. (2007) Characterization of ZNF23, a KRAB-containing protein that is downregulated in human cancers and inhibits cell cycle progression. *Experimental Cell Research*, 313: 254-63.

Huang, D. *et al* (2009) E2-RING expansion of the NEDD8 cascade confers specificity to cullin modification. *Molecular Cell*, 33: 483-95.

Huntley, S., Baggott, D.M., Hamilton, A.T., Tran-Gyamfi, M., Yang, S., Kim, J., Gordon, L., Branscomb, E. and Stubbs, L. (2006) A comprehensive catalog of human KRAB-associated zinc finger genes: insights into the evolutionary history of a large family of transcriptional repressors. *Genome Research*, 16: 669-77.

Iborra, F.J., Pombo, A., Jackson, D.A. and Cook, P.R. (1996) Active polymerases are localized within discrete transcription "factories" in human nuclei. *Journal of Cell Science*, 109: 1427-36.

Iborra, F.J. (2002) The path that RNA takes from the nucleus to the cytoplasm: a trip with some surprises. *Histochemistry and Cell Biology*, 118: 95-103.

Ishov, A.M., Sotnikov, A.G., Negorev, D., Vladimirova, O.V., Neff, N., Kamitani, T., Yeh, E.T., Strauss, J.F.3rd and Maul, G.G. (1999) PML is critical for ND10 formation and recruits the PML-interacting protein daxx to this nuclear structure when modified by SUMO-1. *Journal of Cell Biology*, 147: 221-34.

Ivanov, A.V. *et al* (2007) PHD domain-mediated E3 ligase activity directs intramolecular sumoylation of an adjacent bromodomain required for gene silencing. *Molecular Cell*, 28: 823-37.

Iwase, S., Lan, F., Bayliss, P., de la Torre-Ubieta, L., Huarte, M., Qi, H.H., Whetstine, J.R., Bonni, A., Roberts, T.M. and Shi, Y. (2007) The X-linked mental retardation gene SMCX/JARID1C defines a family of histone H3 lysine 4 demethylases. *Cell*, 128: 1077-88.

Jackson, D.A., Hassan, A.B., Errington, R.J. and Cook, P.R. (1993) Visualization of focal sites of transcription within human nuclei. *EMBO Journal*, 12: 1059-65.

Jackson, D.A., Iborra, F.J., Manders, E.M. and Cook, P.R. (1998) Numbers and organization of RNA polymerases, nascent transcripts, and transcription units in HeLa nuclei. *Molecular and Cellular Biology*, 9: 1523-36.

Jacobson, R.H., Ladurner, A.G., King, D.S. and Tjian, R. (2000) Structure and function of a human TAFII250 double bromodomain module. *Science*, 288: 1422-5.

Jäkel, S. and Görlich, D. (1998) Importin beta, transportin, RanBP5 and RanBP7 mediate nuclear import of ribosomal proteins in mammalian cells. *EMBO Journal*, 17: 4491-502.

Jiang, P.P., Frederick, K., Hansen, T.H. and Miller, R.D. (1996) Localization of the mouse gene releasing sex-limited expression of Slp. *Proceedings of the National Academy of Sciences USA*, 93: 913-7.

Joazeiro, C.A. and Weissman, A.M. (2000) RING finger proteins: mediators of ubiquitin ligase activity. *Cell*, 102: 549-52.

Johnson, D., al-Shawi, R. and Bishop, J.O. (1995) Sexual dimorphism and growth hormone induction of murine pheromone-binding proteins. *Journal of Molecular Endocrinology*, 14: 21-34.

Johnson, E.S. and Blobel, G. (1997) Ubc9p is the conjugating enzyme for the ubiquitin-like protein Smt3p. *Journal of Biological Chemistry*, 272: 26799-802.

Johnson, D.S., Mortazavi, A., Myers, R.M. and Wold, B. (2007) Genome-wide mapping of *in vivo* protein-DNA interactions. *Science*, 316: 1497-502.

Jones, R.S. and Gelbart, W.M. (1993) The *Drosophila* Polycomb-group gene Enhancer of zeste contains a region with sequence similarity to trithorax. *Molecular and Cellular Biology*, 13: 6357-66.

Jones, P.L., Veenstra, G.J., Wade, P.A., Vermaak, D., Kass, S.U., Landsberger, N., Strouboulis, J. and Wolffe, A.P. (1998) Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. *Nature Genetics*, 19: 187-91.

Kaper, J.B., Morris, J.G.Jr and Levine, M.M. (1995) Cholera. *Clinical Microbiology Reviews*, 8: 48-86.

Karras, G.I., Kustatscher, G., Buhecha, H.R., Allen, M.D., Pugieux, C., Sait, F., Bycroft, M. and Ladurner, A.G. (2005) The macro domain is an ADP-ribose binding module. *EMBO Journal*, 24: 1911-20.

Khetchoumian, K., Teletin, M., Mark, M., Lerouge, T., Cerviño, M., Oulad-Abdelghani, M., Chambon, P. and Losson, R. (2004) TIF1_, a novel HP1-interacting member of the transcriptional intermediary factor 1 (TIF1) family expressed by elongating spermatids. *Journal of Biological Chemistry*, 279: 48329-41.

Kießlich, A., von Mikecz, A. and Hemmerich, P. (2002) Cell cycle-dependent association of PML bodies with sites of active transcription in nuclei of mammalian cells. *Journal of Structural Biology*, 140: 167-79.

Kim, S.S., Chen, Y.M., O'Leary, E., Witzgall, R., Vidal, M. and Bonaventre, J.V. (1996) A novel member of the RING finger family, KRIP-1, associates with the KRAB-A transcriptional repressor

domain of zinc finger proteins. *Proceedings of the National Academy of Sciences USA*, 93: 15299-304.

Kim, M.Y., Mauro, S., Gévry, N., Lis, J.T. and Kraus, W.L. (2004) NAD+-dependent modulation of chromatin structure and transcription by nucleosome binding properties of PARP-1. *Cell*, 119: 803-14.

Klose, R.J., Yamane, K., Bae, Y., Zhang, D., Erdjument-Bromage, H., Tempst, P., Wong, J. and Zhang, Y. (2006) The transcriptional repressor JHDM3A demethylates trimethyl histone H3 lysine 9 and lysine 36. *Nature*, 442: 312-6.

Kobza, K., Camporeale, G., Rueckert, B., Kueh, A., Griffin, J.B., Sarath, G. and Zempleni, J. (2005) K4, K9 and K18 in human histone H3 are targets for biotinylation by biotinidase. *FEBS Journal*, 272: 4249-59.

Kosak, S.T., Skok, J.A., Medina, K.L., Riblet, R., Le Beau, M.M., Fisher, A.G. and Singh, H. (2002) Subnuclear compartmentalization of immunoglobulin loci during lymphocyte development. *Science*, 296: 158-62.

Kothapalli, N., Camporeale, G., Kueh, A., Chew, Y.C., Oommen, A.M., Griffin, J.B. and Zempleni, J. (2005) Biological functions of biotinylated histones. *Journal of Nutritional Biochemistry*, 16: 446-8.

Kourmouli, N., Theodoropoulos, P.A., Dialynas, G., Bakou, A., Politou, A.S., Cowell, I.G., Singh, P.B. and Georgatos, S.D. (2000) Dynamic associations of heterochromatin protein 1 with the nuclear envelope. *EMBO Journal*, 19: 6558-68.

Kouzarides, T. (2007) Chromatin modifications and their function. *Cell*, 128: 693-705.

Krebs, C.J., Larkins, L.K., Price, R., Tullis, K.M., Miller, R.D. and Robins, D.M. (2003) Regulator of sex-limitation (Rsl) encodes a pair of KRAB zinc-finger genes that control sexually dimorphic live gene expression. *Genes and Development*, 17: 2664-74.

Kumar, P.P., Bischof, O., Purbey, P.K., Notani, D., Urlaub, H., Dejean, A. and Galande, S. (2007) Functional interaction between PML and SATB1 regulates chromatin-loop architecture and transcription of the MHC class I locus. *Nature Cell Biology*, 9: 45-56.

Kumaran, R.I. and Spector, D.L. (2008) A genetic locus targeted to the nuclear periphery in living cells maintains its transcriptional competence. *Journal of Cell Biology*, 180: 51-65.

Kuo, M.L., Chau, Y.P., Wang, J.H. and Shiah, S.G. (1996) Inhibitors of poly(ADP-ribose) polymerase block nitric oxide-induced apoptosis but not differentiation in human leukemia HL-60 cells. *Biochemical and Biophysical Research Communications*, 219: 502-8.

Kurz, A., Lampel, S., Nickolenko, J.E., Bradl, J., Benner, A., Zirbel, R.M., Cremer, T. And Lichter, P. (1996) Active and inactive genes localize preferentially in the periphery of chromosome territories. *Journal of Cell Biology*, 135: 1195-205.

Kuzmichev, A., Nishioka, K., Erdjument-Bromage, H., Tempst, P. and Reinberg, D. (2002) Histone methyltransferase activity associated with a human multiprotein complex containing the Enhancer of Zeste protein. (2002) *Genes and Development*, 16: 2893-905.

Lachner, M., O'Carroll, D., Rea, S., Mechtler, K. and Jenuwein, T. (2001) Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature*, 410: 116-20.

Lallemand-Breitenbach, V., Zhu, J., Puvion, F., Koken, M., Honoré, N., Doubeikovsky, A., Duprez, E., Pandolfi, P.P., Puvion, E., Freemont, P. and de Thé, H. (2001) Role of promyelocytic leukemia (PML) sumolation in nuclear body formation, 11S proteasome recruitment, and As_2O_3-induced PML or PML/retinoic acid receptor alpha degradation. *Journal of Experimental Medicine*, 193: 1361-71

Lallemand-Breitenbach, V., Jeanne, M., Benhenda, S., Nasr, R., Lei, M., Peres, L., Zhou, J., Zhu, J., Raught, B. and de Thé, H. (2008) Arsenic degrades PML or PML-RAR_ through a SUMO-triggered RNF4/ubiquitin-mediated pathway. *Nature Cell Biology*, 10: 547-55.

Lan, F., Bayliss, P.E., Rinn, J.L., Whetstine, J.R., Wang, J.K., Chen, S., Iwase, S., Alpatov, R., Issaeva, I., Canaani, E., Roberts, T.M., Chang, H.Y. and Shi, Y. (2007) A histone H3 lysine 27 demethylase regulates animal posterior development. *Nature*, 449: 689-94.

Lander, E.S. *et al* (2001) Initial sequencing and analysis of the human genome. *Nature*, 409: 860-921.

Le Douarin, B., Nielsen, A.L., Garnier, J.M., Ichinose, H., Jeanmougin, F., Losson, R. and Chambon, P. (1996) A possible involvement of TIF1_ and TIF1_ in the epigenetic control of transcription by nuclear receptors. *EMBO Journal*, 15: 6701-15.

Lechner, M.S., Begg, G.E., Speicher, D.W. and Rauscher, F.J.3rd (2000) Molecular determinants for targeting heterochromatin protein 1-mediated gene silencing: direct chromoshadow domain-KAP-1 corepressor interaction is essential. *Molecular and Cellular Biology*, 20: 6449-65.

Lee, G.W., Melchior, F., Matunis, M.J., Mahajan, R., Tian, Q. and Anderson, P. (1998) Modification of Ran GTPase-related modifier SUMO-1 requires Ubc9, an E2-type ubiquitin-conjugating enzyme homologue. *Journal of Biological Chemistry*, 273: 6503-7.

Lee, T.I. and Young, R.A. (2000) Transcription of eukaryotic protein-coding genes. *Annual Review of Genetics*, 34: 77-137.

Lee, T.I. *et al* (2006) Control of developmental regulators by Polycomb in human embryonic stem cells. *Cell*, 125: 301-13.

Lee, M.G., Villa, R., Trojer, P., Norman, J., Yan, K.P., Reinberg, D., Di Croce, L. and Shiekhattar, R. (2007a) Demethylation of H3K27 regulates polycomb recruitment and H2A ubiquitination. *Science*, 318: 447-50.

Lee, Y.K., Thomas, S.N., Yang, A.J. and Ann, D.K. (2007b) Doxorubicin down-regulates Krüppel-associated box domain-associated protein 1 sumoylation that relieves its transcription repression on p21WAF1/CIP1 in breast cancer MCF-7 cells. *Journal of Biological Chemistry*, 282: 1595-606.

Lee, S.J., Lee, J.R., Hahn, H.S., Kim, Y.H., Ahn, J.H., Bae, C.D., Yang, J.M. and Hahn, M.J. (2007c) PIAS1 interacts with the KRAB zinc finger protein, ZNF133, via zinc finger motifs and regulates its transcriptional activity. *Experimental and Molecular Medicine*, 39: 450-7.

Lee, J., Thompson, J.R., Botuyan, M.V. and Mer, G. (2008) Distinct binding modes specify the recognition of methylated histones H3K4 and H4K20 by JMJD2A-tudor. *Nature Structural and Molecular Biology*, 15: 109-11.

Lespinet, O., Wolf, Y.I., Koonin, E.V. and Aravind, L. (2002) The role of lineage-specific gene family expansion in the evolution of eukaryotes. *Genome Research*, 12: 1048-59.

Li, Z., Na, X., Wang, D., Schoen, S.R., Messing, E.M. and Wu, G. (2002) Ubiquitination of a novel deubiquitinating enzyme requires direct binding to von Hippel-Lindau tumor suppressor protein. *Journal of Biological Chemistry*, 277: 4656-62.

Li, H., Ilin, S., Wang, W., Duncan, E.M., Wysocka, J., Allis, C.D. and Patel, D.J. (2006) Molecular basis for site-specific read-out of histone H3K4me3 by the BPTF PHD finger of NURF. *Nature*, 442: 91-5.

Li, X., Lee, Y.K., Jeng, J.C., Yen, Y., Schultz, D.C., Shih, H.M. and Ann, D.K. (2007) Role for KAP1 serine 824 phosphorylation and sumoylation/desumoylation switch in regulating KAP1-mediated transcriptional repression. *Journal of Biological Chemistry*, 282: 36177-89.

Lin, D.Y. *et al* (2006) Role of SUMO-interacting motif in Daxx SUMO modification, subnuclear localization, and repression of sumoylated transcription factors. *Molecular Cell*, 24: 341-54.

Litt, M.D., Simpson, M., Gaszner, M., Allis, C.D. and Felsenfeld, G. (2001) Correlation between histone lysine methylation and developmental changes at the chicken _-globin locus. *Science*, 293: 2453-5.

Lockhart, D.J., Dong, H., Byrne, M.C., Follettie, M.T., Gallo, M.V., Chee, M.S., Mittmann, M., Wang, C., Kobayashi, M., Horton, H. and Brown, E.L. (1996) Expression monitoring by hybridisation to high-density oligonucleotide arrays. *Nature Biotechnology*, 14: 1675-80.

Looman, C., Hellman, L. and Abrink, M. (2004) A novel Krüppel-Associated Box identified in a panel of mammalian zinc finger proteins. *Mammalian Genome*, 15: 35-40.

Lorick, K.L., Jensen, J.P., Fang, S., Ong, A.M., Hatakeyama, S. and Weissman, A.M. (1999) RING fingers mediate ubiquitin-conjugating enzyme (E2)-dependent ubiquitination. *Proceedings of the National Academy of Sciences USA*, 96: 11364-9.

Lu, P.J., Sundquist, K., Baeckstrom, D., Poulsom, R., Hanby, A., Meier-Ewert, S., Jones, T., Mitchell, M., Pitha-Rowe, P., Freemont, P. and Taylor-Papadimitriou, J. (1999) A novel gene (PLU-1) containing highly conserved putative DNA/chromatin binding motifs is specifically up-regulated in breast cancer. *Journal of Biological Chemistry*, 274: 15633-45.

Luciani, J.J., Depetris, D., Usson, Y., Metzler-Guillemain, C., Mignon-Ravix, C., Mitchell, M.J., Megarbane, A., Sarda, P., Sirma, H., Moncla, A., Feunteun, J. and Mattei, M.G. (2006) PML nuclear bodies are highly organised DNA-protein structures with a function in heterochromatin remodelling at the G2 phase. *Journal of Cell Science*, 119: 2518-31.

Ludérus, M.E., de Graaf, A., Mattia, E., den Blaauwen, J.L., Grande, M.A., de Jong, L. and van Driel, R. (1992) Binding of matrix attachment regions to lamin B1. *Cell*, 70: 949-59.

Luger, K., Mäder, A.W., Richmond, R.K., Sargent, D.F. and Richmond, T.J. (1997) Crystal structure of the nucleosome core particle at 2.8Å resolution. *Nature*, 389: 251-60.

Mahajan, R., Delphin, C., Guan, T., Gerace, L. and Melchior, F. (1997) A small ubiquitin-related polypeptide involved in targeting RanGAP1 to nuclear pore complex protein RanBP2. *Cell*, 88: 97-107.

Mahy, N.L., Perry, P.E., Gilchrist, S., Baldock, R.A. and Bickmore, W.A. (2002) Spatial organization of active and inactive genes and noncoding DNA within chromosome territories. *Journal of Cell Biology*, 157: 579-89.

Mahy, N.L., Perry, P.E. and Bickmore, W.A. (2002b) Gene density and transcription influence the localization of chromatin outside of chromosome territories detectable by FISH. *Journal of Cell Biology*, 159: 753-63.

Makowski, G.S. and Sunderman, F.W.Jr. (1992) The interactions of zinc, nickel, and cadmium with *Xenopus* transcription factor IIIA, assessed by equilibrium dialysis. *Journal of Inorganic Biochemistry*, 48: 107-19.

Margolin, J.F., Friedman, J.R., Meyer, W.K., Vissing, H., Thiessen, H.J. and Rauscher, F.J.3rd (1994) Krüppel-associated boxes are potent transcriptional repression domains. *Proceedings of the National Academy of Sciences USA*, 91: 4509-13.

Margueron, R., Li, G., Sarma, K., Blais, A., Zavadil, J., Woodcock, C.L., Dynlacht, B.D. and Reinberg, D. (2008) Ezh1 and Ezh2 maintain repressive chromatin through different mechanisms. *Molecular Cell*, 32: 503-18.

Mark, C., Abrink, M., Hellman, L. (1999) Comparative analysis of KRAB zinc finger proteins in rodents and man: evidence for several evolutionarily distinct subfamilies of KRAB zinc finger genes. *DNA Cell Biology*, 18: 381-96.

Martens, J.A. and Winston, F. (2003) Recent advances in understanding chromatin remodelling by Swi/Snf complexes. *Current Opinion in Genetics and Development*, 13: 136-42.

Martin, S., Nishimune, A., Mellor, J.R. and Henley, J.M. (2007) SUMOylation regulates kainite-receptor-mediated synaptic transmission. *Nature*, 447: 321-5.

Mascle, X.H., Germain-Desprez, D., Huynh, P., Estephan, P. and Aubry, M. (2007) Sumoylation of the transcriptional intermediary factor 1_ (TIF1_), the co-repressor of the KRAB multifinger proteins, is required for its transcriptional activity and is modulated by the KRAB domain. *Journal of Biological Chemistry*, 282: 10190-202.

Masson, M., Niedergang, C., Schreiber, V., Muller, S., Menissier-de Murcia, J. and de Murcia, G. (1998) XRCC1 is specifically associated with poly(ADP-ribose) polymerase and negatively regulates its activity following DNA damage. *Molecular and Cellular Biology*, 18: 3563-71.

Matsuda, E., Agata, Y., Sugai, M., Katakai, T., Gonda, H. and Shimizu, A. (2001) Targeting of Krüppel-associated box-containing zinc finger proteins to centromeric heterochromatin. Implication for the gene silencing mechanisms. *Journal of Biological Chemistry*, 276: 14222-9.

Matunis, M.J., Coutavas, E. and Blobel, G. (1996) A novel ubiquitin-like modification modulates the partitioning of the Ran-GTPase-activating protein RanGAP1 between the cytosol and the nuclear pore complex. *Journal of Cell Biology*, 135: 1457-70.

Maul, G.G., Negorev, D., Bell, P. and Ishov, A.M. (2000) Review: properties and assembly mechanisms of ND10, PML bodies, or PODs. *Journal of Structural Biology*, 129: 278-87.

Meehan, R.R., Lewid, J.D., McKay, S., Kleiner, E.L. and Bird, A.P. (1989) Identification of a mammalian protein that binds specifically to DNA containing methylated CpGs. *Cell*, 58: 499-507.

Meehan, R.R. (2003) DNA methylation in animal development. *Seminars in Cell and Developmental Biology*, 14: 53-65.

Meissner, A., Mikkelsen, T.S., Gu, H., Wernig, M., Hanna, J., Sivachenko, A., Zhang, X., Bernstein, B.E., Nusbaum, C., Jaffe, D.B., Gnirke, A., Jaenisch, R. and Lander, E.S. (2008) Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature*, 454: 766-70.

Meroni, G. and Diez-Roux, G. (2005) TRIM/RBCC, a novel class of 'single protein RING finger' E3 ubiquitin ligases. *Bioessays*, 27: 1147-57.

Mikkelsen, T.S. *et al* (2007) Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature*, 448: 553-60.

Milne, T.A., Briggs, S.D., Brock, H.W., Martin, M.E., Gibbs, D., Allis, C.D. and Hess, J.L. (2002) MLL targets SET domain methyltransferase activity to Hox gene promoters. *Molecular Cell*, 10: 1107-17.

Minc, E., Allory, Y., Worman, H.J., Courvalin, J.C. and Buendia, B. (1999) Localization and phosphorylation of HP1 proteins during the cell cycle in mammalian cells. *Chromosoma*, 108: 220-34.

Minc, E., Courvalin, J.C. and Buendia, B. (2000) HP1_ associates with euchromatin and heterochromatin in mammalian nuclei and chromosomes. *Cytogenetics and Cell Genetics*, 90: 279-84.

Minc, E., Allory, Y., Courvalin, J.C. and Buendia, B. (2001) Immunolocalization of HP1 proteins in metaphasic mammalian chromosomes. *Methods in Cell Science*, 23: 171-4.

Moen, P.T.Jr., Johnson, C.V., Byron, M., Shopland, L.S., de la Serna, I.L., Imbalzano, A.N. and Lawrence, J.B. (2004) Repositioning of muscle-specific genes relative to the periphery of SC-35 domains during skeletal myogenesis. *Molecular and Cellular Biology*, 15: 197-206.

Montgomery, N.D., Yee, D., Chen, A., Kalantry, S., Chamberlain, S.J., Otte, A.P. and Magnuson, T. (2005) The murine polycomb group protein Eed is required for global histone H3 lysine-27 methylation. *Current Biology*, 15: 942-7.

Moosmann, P., Georgiev, O., Le Douarin, B., Bourquin, J.P. and Schaffner, W. (1996) Transcriptional repression by RING finger protein TIF1_ that interacts with the KRAB repressor domain of KOX1. *Nucleic Acids Research*, 24: 4859-67.

Morey, C., Da Silva, N.R., Perry, P. and Bickmore, W.A. (2007) Nuclear reorganisation and chromatin decondensation are conserved, but distinct, mechanisms linked to Hox gene activation. *Development*, 134: 909-19.

Morey, C., Da Silva, N.R., Kmita, M., Duboule, D. and Bickmore, W.A. (2008) Ectopic nuclear reorganisation driven by a Hoxb1 transgene transposed into Hoxd. *Journal of Cell Science*, 121: 571-7.

Mouchantaf, R., Azakir, B.A., McPherson, P.S., Millard, S.M., Wood, S.A. and Angers, A. (2006) The ubiquitin ligase Itch is auto-ubiquitylated *in vivo* and *in vitro* but is protected from degradation by interacting with the deubiquitylating enzyme FAM/USP9X. *Journal of Biological Chemistry*, 281: 38738-47.

Müller, J., Hart, C.M., Francis, N.J., Vargas, M.L., Sengupta, A., Wild, B., Miller, E.L., O'Connor, M.B., Kingston, R.E. and Simon, J.A. (2002) Histone methyltransferase activity of a Drosophila Polycomb group repressor complex. *Cell*, 111: 197-208.

Nacerddine, K., Lehembre, F., Bhaumik, M., Artus, J., Cohen-Tannoudji, M., Babinet, C., Pandolfi, P.P. and Dejean, A. (2005) The SUMO pathway is essential for nuclear integrity and chromosome segregation in mice. *Developmental Cell*, 9: 769-779.

Nan, X., Ng, H.H., Johnson, C.A., Laherty, C.D., Turner, B.M., Eisenman, R.N. and Bird, A. (1998) Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. *Nature*, 393: 386-9.

Narang, M.A., Dumas, R., Ayer, L.M. and Gravel, R.A. (2004) Reduced histone biotinylation in multiple carboxylase deficiency patients: a nuclear role for holocarboxylase synthetase. *Human Molecular Genetics*, 13: 15-23.

Nathan, D., Ingvarsdottir, K., Sterner, D.E., Bylebyl, G.R., Dokmanovic, M., Dorsey, J.A., Whelan, K.A., Krsmanovic, M., Lane, W.S., Meluh, P.B., Johnson, E.S. and Berger, S.L. (2006) Histone sumoylation is a negative regulator in *Saccharomyces cerevisiae* and shows dynamic interplay with positive-acting histone modifications. *Genes and Development*, 20: 966-76.

Negorev, D. and Maul, G.G. (2001) Cellular protein localized at and interacting within ND10/PML nuclear bodies/PODs suggest functions of a nuclear depot. *Oncogene*, 20: 7234-42.

Ng, H.H., Zhang, Y., Hendrich, B., Johnson, C.A., Turner, B.M., Erdjument-Bromage, H., Tempst, P., Reinberg, D. and Bird, A. (1999) MBD2 is a transcriptional repressor belonging to the MeCP1 histone deacetylase complex. *Nature Genetics*, 23: 58-61.

Nielsen, A.L., Ortiz, J.A., You, J., Oulad-Abdelghani, M., Khechumian, R., Gansmuller, A., Chambon, P. and Losson, R. (1999) Interaction with members of the heterochromatin protein 1 (HP1) family and histone deacetylation are differentially involved in transcriptional silencing by members of the TIF1 family. *EMBO Journal*, 18: 6385-95.

Nielsen, A.L., Jørgensen, P., Lerouge, T., Cerviño, M., Chambon, P. And Losson, R. (2004) Nizp1, a novel multitype zinc finger protein that interacts with the NSD1 histone lysine methyltransferase through a unique C2HR motif. *Molecular and Cellular Biology*, 24: 5184-96.

Nishioka, K., Chuikov, S., Sarma, K., Erdjument-Bromage, H., Allis, C.D., Tempst, P. and Reinberg, D. (2002) Set9, a novel histone H3 methyltransferase that facilitates transcription by precluding histone tail modification required for heterochromatin formation. *Genes and Development*, 16: 479-89.

Noma, K., Allis, C.D. and Grewal, S.I. (2001) Transitions in distinct histone H3 methylation patterns at the heterochromatin domain boundaries. *Science*, 293: 1150-5.

Nowak, S.J. and Corces, V.G. (2004) Phosphorylation of histone H3: a balancing act between chromosome condensation and transcriptional activation. *Trends in Genetics*, 20: 214-20.

Nyman, U., Hallman, H., Hadlaczky, G., Pettersson, I., Sharp, G. and RIngertz, N.R. (1986) Intranuclear localization of snRNP antigens. *Journal of Cell Biology*, 102: 137-44.

O'Carroll, D. *et al* (2000) Isolation and characterization of Suv39h2, a second histone H3 methyltransferase gene that displays testis-specific expression. *Molecular and Cellular Biology*, 20: 9423-33.

O'Carroll, D., Erhardt, S., Pagani, M., Barton, S.C., Surani, M.A. and Jenuwein, T. (2001) The polycomb-group gene Ezh2 is required for early mouse development. *Molecular and Cellular Biology*, 21: 4330-6.

O'Geen, H., Squazzo, S.L., Iyengar, S., Blahnik, K., Rinn, J.L., Chang, H.Y., Green, R. and Farnham, P. (2007) Genome-wide analysis of KAP1 binding suggests autoregulation of KRAB-ZNFs. *PLoS Genetics*, 3: e89.

Okano, M., Bell, D.W., Haber, D.A. and Li, E. (1999) DNA methyltransferases Dnmt3a and Dnmt3b are essential for *de novo* methylation and mammalian development. *Cell*, 99:247-57.

Ong, S.E., Blagoev, B., Kratchmarova, I., Kristensen, D.B., Steen, H., Pandey, A. and Mann, M. (2002) Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Molecular and Cellular Proteomics*, 1: 376-86.

Orphanides, G., LeRoy, G., Chang, C.H., Luse, D.S. and Reinberg, D. (1998) FACT, a factor that facilitates transcript elongation through nucleosomes. *Cell*, 92: 105-16.

Osborne, C.S., Chakalova, L., Brown, K.E., Carter, D., Horton, A., Debrand, E., Goyenechea, B., Mitchell, J.A., Lopes, S., Reik, W. and Fraser, P. (2004) Active genes dynamically colocalize to shared sites of ongoing transcription. *Nature Genetics*, 36: 1065-71.

Ou, Y., Wang, S., Cai, Z., Wang, Y., Wang, C., Li, Y., Li, F., Yuan, W., Liu, B., Wu, X. and Liu, M. (2005) ZNF328, a novel human zinc-finger protein, suppresses transcriptional activities of SRE and AP-1. *Biochemical and Biophysical Research Communications*, 333: 1034-44.

Owen, D.J., Ornaghi, P., Yang, J.C., Lowe, N., Evans, P.R., Ballario, P., Neuhaus, D., Filetici, P. and Travers, A.A. (2000) The structural basis for the recognition of acetylated histone H4 by the bromodomain of histone acetyltransferase gcn5p. *EMBO Journal*, 19: 6141-9.

Palstra, R.J., Tolhuis, B., Splinter, E., Nijmeijer, R., Grosveld, F. and de Laat, W. (2003) The _-globin nuclear compartment in development and erythroid differentiation. *Nature Genetics*, 35: 190-4.

Palstra, R.J., Simonis, M., Klous, P., Brasset, E., Eijkelkamp, B. and de Laat, W. (2008) Maintenance of long-range DNA interactions after inhibition of ongoing RNA polymerase II transcription. *PLoS ONE*, 3: e1661.

Pardo, P.S., Leung, J.K., Lucchesi, J.C. and Pereira-Smith, O.M. (2002) MRG15, a novel chromodomain protein, is present in two distinct multiprotein complexes involved in transcriptional activation. *Journal of Biological Chemistry*, 277: 50860-6.

Pasini, D., Bracken, A.P., Jensen, M.R., Lazzerini Denchi, E. and Helin, K. (2004) Suz12 is essential for mouse development and for EZH2 histone methyltransferase activity. *EMBO Journal*, 23: 4061-71.

Pavri, R., Lewis, B., Kim, T.K., Dilworth, F.J., Erdjument-Bromage, H., Tempst, P., de Murcia, G., Evans, R., Chambon, P. and Reinberg, D. (2005) PARP-1 determines specificity in a retinoid signalling pathway via direct modulation of mediator. *Molecular Cell*, 18: 83-96.

Payen, E., Verkerk, T., Michalovich, D., Dreyer, S.D., Winterpacht, A., Lee, B., De Zeeuw, C.I., Grosveld, F. and Galjart, N. (1998) The centromeric/nucleolar chromatin protein ZFP-37 may function to specify neuronal nuclear domains. *Journal of Biological Chemistry*, 273: 9099-109.

Payre, F., Noselli, S., Lefrére, V. and Vincent, A. (1990) The closely related *Drosophila* sry _ and sry _ zinc finger proteins show differential embryonic expression and distinct patterns of binding sites on polytene chromosomes. *Development*, 110: 141-9.

Peña, P.V., Davrazou, F., Shi, X., Walter, K.L., Verkhusha, VV., Gozani, O., Zhao, R. and Kutateladze, T.G. (2006) Molecular mechanism of histone H3K4me3 recognition by plant homeodomain of ING2. *Nature*, 442: 100-3.

Peng, H., Feldman, I. and Rauscher, F.J.3rd (2002) Hetero-oligomerisation among the TIF family of RBCC/TRIM domain-containing nuclear cofactors: a potential mechanism for regulating the switch between coactivation and corepression. *Journal of Molecular Biology*, 320: 629-44.

Pengue, G., Calabrò, V., Bartoli, P.C., Pagliuca, A. and Lania, L. (1994) Repression of transcriptional activity at a distance by the evolutionarily conserved KRAB domain present in a subfamily of zinc finger proteins. *Nucleic Acids Research*, 22: 2908-14.

Peters, D.M., Griffin, J.B., Stanley, J.S., Beck, M.M. and Zempleni, J. (2002) Exposure to UV light causes increased biotinylation of histones in Jurkat cells. *American Journal of Physiology – Cell Physiology*, 283: C878-84.

Pokholok, D.K. *et al* (2005) Genome-wide map of nucleosome acetylation and methylation in yeast. *Cell*, 122: 517-27.

Pott, P.R. and Yu, H. (2007) The SMC5/6 complex maintains telomere length in ALT cancer cells through SUMOylation of telomere-binding proteins. *Nature Structural and Molecular Biology*, 14: 581-90.

Prokhortchouk, A., Hendrich, B., Jørgensen, H., Ruzov, A., Wilm, M., Georgiev, G., Bird, A. and Prokhortchouk, E. (2001) The p21 catenin partner Kaiso is a DNA methylation-dependent transcriptional repressor. *Genes and Development*, 15: 1613-8.

Purnell, M.R. and Whish, W.J. (1980) Novel inhibitors of poly(ADP-ribose) synthetase. *Biochemical Journal*, 185: 775-7.

Qi, X., Li, Y., Xiao, J., Yuan, W., Yan, Y., Wang, Y., Liang, S., Zhu, C., Chen, Y., Liu, M. and Wu, X. (2006) Activation of transcriptional activities of AP-1 and SRE by a new zinc-finger protein ZNF641. *Biochemical and Biophysical Research Communications*, 339: 1155-64.

Quénet, D., Gasser, V., Fouillen, L., Cammas, F., Sanglier-Cianferani, S., Losson, R. and Dantzer, F. (2008) The histone subcode: poly(ADP-ribose) polymerase-1 (Parp-1) and Parp-2 control cell differentiation by regulating the transcriptional intermediary factor TIF1_ and the heterochromatin protein HP1_. *FASEB Journal*, 22: 3853-65.

Racki, L.R. and Narlikar, G.J. (2008) ATP-dependent chromatin remodelling enzymes: two heads are not better, just different. *Current Opinion in Genetics and Development*, 18: 137-44.

Rajan, S., Plant, L.D., Rabin, M.L., Butler, M.H. and Goldstein, S.A. (2005) Sumoylation silences the plasma membrane leak K+ channel K2P1. *Cell*, 121: 37-47.

Rajesh, M., Mukhopadhyay, P., Bátkai, S., Godlewski, G., Haskó, G., Liaudet, L. and Pacher, P. (2006) Pharmocological inhibition of poly(ADP-ribose) polymerase inhibits angiogenesis. *Biochemical Biophysics Research Communications*, 350: 352-7.

Rea, S., Eisenhaber, F., O'Carroll, D., Strahl, B.D., Sun, Z.W., Schmid, M., Opravil, S., Mechtler, K., Ponting, C.P., Allis, C.D. and Jenuwein, T. (2000) Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature*, 406: 593-9.

Reddy, K.L., Zullo, J.M., Bertolino, E. and Singh, H. (2008) Transcriptional repression mediated by repositioning of genes to the nuclear lamina. *Nature*, 452: 243-7.

Remboutsika, E., Lutz, Y., Gansmuller, A., Vonesch, J.L., Losson, R. and Chambon, P. (1999) The putative nuclear receptor mediator TIF1_ is tightly associated with euchromatin. *Journal of Cell Science*, 112: 1671-83.

Remboutsika, E., Yamamoto, K., Harbers, M. and Schmutz, M. (2002) The bromodomain mediates transcriptional intermediary factor 1_-nucleosome interactions. *Journal of Biological Chemistry*, 277: 50318-25.

Reverter, D. and Lima, C.D. (2005) Insights into E3 ligase activity revealed by a SUMO-RanGAP1-Ubc9-Nup358 complex. *Nature*, 435: 687-92.

Reymond, A. *et al* (2001) The tripartite motif family identifies cell compartments. *EMBO Journal*, 20: 2140-51.

Rice, J.C., Briggs, S.D., Ueberheide, B., Barber, C.M., Shabanowitz, J., Hunt, D.F., Shinkai, Y. and Allis, C.D. (2003) Histone methyltransferases direct different degrees of methylation to define distinct chromatin domains. *Molecular Cell*, 12: 1591-8.

Robinson, P.J., Fairall, L., Huynh, V.A. and Rhodes, D. (2006) EM measurements define that dimensions of the "30-nm" chromatin fiber: evidence for a compact, interdigitated structure. *Proceedings of the National Academy of Sciences USA*, 103: 6506-11.

Rodriguez, M.S., Dargemont, C. and Hay, R.T. (2001) SUMO-1 conjugation *in vivo* requires both a consensus modification motif and nuclear targeting. *Journal of Biological Chemistry*, 276: 12654-9.

Roesch, A., Mueller, A.M., Stempfl, T., Moehle, C., Landthaler, M. and Vogt, T. (2008) RBP2-H1/JARID1B is a transcriptional regulator with a tumor suppressive potential in melanoma cells. *International Journal of Cancer*, 122: 1047-57.

Roque, A., Ponte, I., Arrondo, J.L. and Suau, P. (2008) Phosphorylation of the carboxy-terminal domain of histone H1: effects on secondary structure and DNA condensation. *Nucleic Acids Research*, 36: 4719-26.

Rual, J.F. *et al* (2006). Towards a proteome-scale map of the human protein-protein interaction network. *Nature*, 437: 1173-8.

Rubin, G.M. *et al* (2000) Comparative genomics of the eukaryotes. *Science*, 287: 2204-15.

Ruiz i Altaba, A., Perry-O'Keefe, H. and Melton, D.A. (1987) *Xfin*: an embryonic gene encoding a multifingered protein in *Xenopus*. *EMBO Journal*, 6: 3065-70.

Ryan, R.F., Schultz, D.C., Ayyanathan, K., Singh, P.B., Friedman, J.R., Fredericks, W.J. and Rauscher, F.J.3rd (1999) KAP-1 corepressor protein interacts and colocalizes with heterochromatic and euchromatic HP1 proteins: a potential role for Krüppel-associated box-zinc finger proteins in heterochromatin-mediated gene silencing.

Sadoni, N., Langer, S., Fauth, C., Bernardi, G., Cremer, T., Turner, B.M. and Zink, D. (1999) Nuclear organization of mammalian genomes. Polar chromosome territories build up functionally distinct higher order compartments. *Journal of Cell Biology*, 146: 1211-26.

Saha, A., Wittmeyer, J. and Cairns, B.R. (2006) Chromatin remodelling: the industrial revolution of DNA around histones. *Nature Reviews Molecular Cell Biology*, 7: 437-47.

Saitoh, H., Pizzi, M.D. and Wang, J. (2002) Peturbation of SUMOlation enzyme Ubc9 by distinct domain within nucleoporin RanBP2/Nup358. *Journal of Biological Chemistry*, 277: 4755-63.

Sala, A., La Rocca, G., Burgio, G., Kotova, E., Di Gesù, D., Collesano, M., Ingrassia, A.M., Tulin, A.V. and Corona, D.F. (2008) The nucleosome-remodeling ATPase ISWI is regulated by poly-ADP-ribosylation. *PLoS Biology*, 6: e252.

Salomoni, P., Ferguson, B.J., Wyllie, A.H. and Rich, T. (2008) New insights into the role of PML in tumour suppression. *Cell Research*, 18: 622-40.

Sampson, D.A., Wang, M. and Matunis, M.J. (2001) The small ubiquitin-like modifier-1 (SUMO-1) consensus sequence mediates Ubc9 binding and is essential for SUMO-1 modification. *Journal of Biological Chemistry*, 276: 21664-9.

Sarraf, S.A. and Stancheva, I. (2004) Methyl-CpG binding protein MBD1 couples histone H3 methylation at lysine 9 by SETDB1 to DNA replication and chromatin assembly. *Molecular Cell*, 15: 595-605.

Saunders, W.S., Chue, C., Goebl, M., Craig, C., Clark, R.F., Powers, J.A., Eissenberg, J.C., Elgin, S.C., Rothfield, N.F. and Earnshaw, W.C. (1993) Molecular cloning of a human homologue of *Drosophila* heterochromatin protein HP1 using anti-centromere autoantibodies with anti-chromo specificity. *Journal of Cell Science*, 104: 573-82.

Scheffner, M., Huibregtse, J.M., Vierstra, R.D. and Howley, P.M. (1993) The HPV-16 E6 and E6-AP complex functions as a ubiquitin-protein ligase in the ubiquitination of p53. *Cell*, 75: 495-505.

Schreiber, V., Amé, J.C., Dollé, P., Schultz, I., Rinaldi, B., Fraulob, V., Ménissier-de Murcia, J. and de Murcia, G. (2002) Poly(ADP-ribose) polymerase-2 (PARP-2) is required for efficient base excision DNA repair in association with PARP-1 and XRCC1. *Journal of Biological Chemistry*, 277: 23028-36.

Schreiber, V., Dantzer, F., Amé, J.C. and de Murcia, G. (2006) Poly(ADP-ribose): novel functions for an old molecule. *Nature Reviews Molecular Cell Biology*, 7: 517-28.

Schultz, D.C., Friedman, J.R. and Rauscher, F.J.3rd (2001) Targeting histone deacetylase complexes via KRAB-zinc finger proteins: the PHD and bromodomains of KAP-1 form a cooperative unit that recruits a novel isoform of the Mi-2_ subunit of NuRD. *Genes and Development*, 15: 428-43.

Schultz, D.C., Ayyanathan, K., Negorev, D., Maul, G.G. and Rauscher, F.J.3rd (2002) SETDB1: a novel KAP-1-associated histone H3, lysine 9-specific methyltransferase that contributes to Hp1-mediated silencing of euchromatic genes by KRAB zinc-finger proteins. *Genes and Development*, 16: 919-32.

Schulze, W.X. and Mann, M. (2004) A novel proteomic screen for peptide-protein interactions. *Journal of Biological Chemistry*, 279: 10756-64.

Schüpbach, T. And Wieschaus, E. (1989) Female sterile mutations on the second chromosome of *Drosophila melanogaster*. II. Mutations blocking oogenesis or altering egg morphology. *Genetics*, 129: 1119-36.

Seeler, J.S., Marchio, A., Sitterlin, D., Transy, C. and Dejean, A. (1998) Interaction of SP100 with HP1 proteins: a link between the promyelocytic leukemia-associated nuclear bodies and the chromatin compartment. *Proceedings of the National Academy of Sciences USA*, 95: 7316-21.

Seeler, J.S., Marchio, A., Losson, R., Desterro, J.M., Hay, R.T., Chambon, P. and Dejean, A. (2001) Common properties of nuclear body protein SP100 and TIF1_ chromatin factor: role of SUMO modification. *Molecular and Cellular Biology*, 21: 3314-24.

Seward, D.J., Cubberley, G., Kim, S., Schonewald, M., Zhang, L., Tripet, B. and Bentley, D.L. (2007) Demethylation of trimethylated histone H3 Lys4 *in vivo* by JARID1 JmjC proteins. *Nature Structural and Molecular Biology*, 14: 240-2.

Shannon, M., Hamilton, A.T., Gordon, L., Branscomb, E. and Stubbs, L. (2003) Differential expansion of zinc-finger transcription factor loci in homologous human and mouse gene clusters. *Genome Research*, 13: 1097-110.

Sharif, J. *et al* (2007) The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA. *Nature*, 450: 908-12.

Shema, E. *et al* (2008) The histone H2B-specific ubiquitin ligase RNF20/hBRE1 acts as a putative tumor suppressor through selective regulation of gene expression. *Genes and Development*, 22: 2664-76.

Shen, T.H., Lin, H.K., Scaglioni, P.P., Yung, T.M. and Pandolfi, P.P. (2006) The mechanisms of PML-nuclear body formation. *Molecular Cell*, 24: 331-9.

Shen, X., Liu, Y., Hsu, Y.J., Fujiwara, Y., Kim, J., Mao, X., Yuan, G.C. and Orkin, S.H. (2008) EZH1 mediates methylation on histone H3 lysine 27 and complements EZH2 in maintaining stem cell identity and executing pluripotency. *Molecular Cell*, 32: 491-502.

Shi, Y., Lan, F., Matson, C., Mulligan, P., Whetstine, J.R., Cole, P.A., Casero, R.A. and Shi, Y. (2004) Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell*, 119: 941-53.

Shiels, C., Islam, S.A., Vatcheva, R., Sasieni, P., Sternberg, M.J., Freemont, P.S. and Sheer, D. (2001) PML bodies associate specifically with the MHC gene cluster in interphase nuclei. *Journal of Cell Science*, 114: 3705-16.

Shiio, Y. and Eisenman, R.N. (2003) Histone sumoylation is associated with transcriptional repression. *Proceedings of the National Academy of Sciences USA*, 100: 13225-30.

Shopland, L.S., Johnson, C.V., Byron, M., McNeil, J. and Lawrence, J.B. (2003) Clustering of multiple specific genes and gene-rich R-bands around SC-35 domains: evidence for local euchromatic neighbourhoods. *Journal of Cell Biology*, 162: 981-90.

Simonis, M., Klous, P., Splinter, E., Moshkin, Y., Willemsen, R., de Wit, E., van Steensel, B. and de Laat, W. (2006) Nuclear organization of active and inactive chromatin domains uuncovered by chromosome conformation capture-on-chip (4C). *Nature Genetics*, 38: 1348-54.

Singh, P.B., Miller, J.R., Pearce, J., Kothary, R., Burton, R.D., Paro,, R., James, T.C. and Gaunt, S.J. (1991) A sequence motif found in a *Drosophila* heterochromatin protein is conserved in animals and plants. *Nucleic Acids Research*, 19: 789-94.

Skapek, S.X., Jansen, D., Wei, T.F., McDermott, T., Huang, W., Olson, E.N. and Lee, E.Y. (2000) Cloning and characterization of a novel Krüppel-associated box family transcriptional repressor that interacts with the retinoblastoma gene product, RB. *Journal of Biological Chemistry*, 275: 7212-23.

Smith, S. (2001) The world according to PARP. *Trends in Biochemical Sciences*, 26: 174-9.

Smothers, J.F. and Henikoff, S. (2001) The hinge and chromo shadow domain impart distinct targeting of HP1-like proteins. *Molecular and Cellular Biology*, 21: 2555-69.

Somech, R., Shaklai, S., Geller, O., Amariglio, N., Simon, A.J., Rechavi, G. and Gal-Yam, E.N. (2005) The nuclear-envelope protein and transcriptional repressor LAP2_ interacts with HDAC3 at the nuclear periphery, and induces histone H4 deacetylation. *Journal of Cell Science*, 118: 4017-25.

Spector, D.L., Fu, X.D. and Maniatis, T. (1991) Associations between distinct pre-mRNA splicing components and the cell nucleus. *EMBO Journal*, 10: 3467-81.

Spector, D.L. (1993) Macromolecular domains within the cell nucleus. *Annual Review of Cell Biology*, 9: 265-315.

Spector, D.L. (2001) Nuclear domains. *Journal of Cell Science*, 114: 2891-3.

Splinter, E., Heath, H., Kooren, J., Palstra, R.J., Klous, P., Grosveld, F., Galjart, N. and de Laat, W. (2006) CTCF mediates long-range chromatin looping and local histone modification in the _-globin locus. *Genes and Development*, 20: 2349-54.

Sproul, D., Gilbert, N. and Bickmore, W.A. (2005) The role of chromatin structure in regulating the expression of clustered genes. *Nature Reviews Genetics*, 6: 775-81.

Sripathy, S.P., Stevens, J. and Schultz, D.C. (2006) The KAP1 corepressor functions to coordinate the assembly of *de novo* HP1-demarcated microenvironments of heterochromatin required for KRAB zinc finger protein-mediated transcriptional repression. *Molecular and Cellular Biology*, 26: 8623-38.

Stade, K., Vogel, F., Schwienhorst, I., Meusser, B., Volkwein, C., Nentwig, B., Dohmen, R.J. and Sommer, T. (2002) A lack of SUMO conjugation affects cNLS-dependent nuclear protein import in yeast. *Journal of Biological Chemistry*, 277: 49554-61.

Stielow, B., Sapetschnig, A., Krüger, I., Kunert, N., Brehm, A., Boutros, M. and Suske, G. (2008) Identification of SUMO-dependent chromatin-associated transcriptional repression components by a genome-wide RNAi screen. *Molecular Cell*, 29: 742-54.

Strahl, B.D. and Allis, C.D. (2000) The language of covalent histone modifications. *Nature*, 403: 41-5.

Sueyoshi, T., Yokomori, N., Korach, K.S. and Negishi, M. (1999) Developmental action of estrogen receptor-alpha feminizes the growth hormone-Stat5b pathway and expression of Cyp2a4 and Cyp2d9 genes in mouse liver. *Molecular Pharmocology*, 56: 473-7.

Sun, H.B., Shen, J. and Yokota, H. (2000) Size-dependent positioning of human chromosomes in interphase nuclei. *Biophysical Journal*, 79: 184-90.

Sutherland, H.G., Mumford, G.K., Newton, K., Ford, L.V., Farrall, R., Dellaire, G., Cáceres, J.F. and Bickmore, W.A. (2001) Large-scale identification of mammalian proteins localized to nuclear sub-compartments. *Human Molecular Genetics*, 10: 1995-2011.

Tachibana, M., Sugimoto, K., Fukushima, T. And Shinkai, Y. (2001) Set domain-containing protein, G9a, is a novel lysine-preferring mammalian histone methyltransferase with hyperactivity and specific selectivity to lysines 9 and 27 of histone H3. *Journal of Biological Chemistry*, 276: 25309-17.

Tan, W., Zheng, L., Lee, W.H. and Boyer, T.G. (2004) Functional dissection of transcription factor ZBRK1 reveals zinc fingers with dual roles in DNA-binding and BRCA1-dependent transcriptional repression. *Journal of Biological Chemistry*, 279: 6576-87.

Tan, W., Kim, S. and Boyer, T.G. (2004b) Tetrameric oligomerization mediates transcriptional repression by the BRCA1-dependent Krüppel-associated box-zinc finger protein ZBRK1. *Journal of Biological Chemistry*, 279: 55153-60.

Tanaka, K., Tsumaki, N., Kozak, C.A., Matsumoto, Y., Nakatani, F., Iwamoto, Y. and Yamada, Y. (2002) A Krüppel-associated box-zinc finger protein, NT2, represses cell-type-specific promoter activity of the alpha 2 (XI) collagen gene. *Molecular and Cellular Biology*, 22: 4256-67.

Tang, J., Chang, H.Y. and Yang, X. (2005) The death domain-associated protein modulates activity of the transcription co-factor Skip/NcoA62. *FEBS Letters*, 579: 2883-90.

Taniura, H., Glass, C. and Gerace, L. (1995) A chromatin binding site in the tail domain of nuclear lamins that interacts with core histones. *Journal of Cell Biology*, 131: 33-44.

Tassan, J.P., Jaquenoud, M., Fry, A.M., Frutiger, S., Hughes, G.J. and Nigg, E.A. (1995) In vitro assembly of a functional human CDK7-cyclin H complex requires MAT1, a novel 36 kDa RING finger protein. *EMBO Journal*, 14: 5608-17.

Tatham, M.H., Geoffroy, M.C., Shen, L., Plechanovova, A., Hattersley, N., Jaffray, E.G., Palvimo, J.J. and Hay, R.T. (2008) RNF4 is a poly-SUMO-specific E3 ubiquitin ligase required for arsenic-induced PML degradation. *Nature Cell Biology*, 10: 538-46.

Taverna, S.D., Ueberheide, B.M., Liu, Y., Tackett, A.J., Diaz, R.L., Shabanowitz, J., Chait, B.T., Hunt, D.F. and Allis, C.D. (2007) Long-distance combinatorial linkage between methylation and acetylation on histone H3 N termini. *Proceedings of the National Academy of Sciences USA*, 104: 2086-91.

Thiriet, C. And Hayes, J.J. (2005) Chromatin in need of a fix: phosphorylation of H2AX connects chromatin to DNA repair. *Molecular Cell*, 18: 617-22.

Thomson, S., Clayton, A.L. and Mahadevan, L.C. (2001) Independent dynamic regulation of histone phosphorylation and acetylation during immediate-early gene induction. *Molecular Cell*, 8: 1231-41.

Tolhuis, B., Palstra, R.J., Splinter, E., Grosveld, F. and de Laat, W. (2002) Looping and interaction between hypersensitive sites in the active _-globin locus. *Molecular Cell*, 10: 1453-65.

Tominaga, K., Kirtane, B., Jackson, J.G., Ikeno, Y., Ikeda, T., Hawks, C., Smith, J.R., Matzuk, M.M. and Pereira-Smith, O.M. (2005) *Molecular and Cellular Biology*, 25: 2924-37.

Torres-Padilla, M.E. and Zernicka-Goetz, M. (2006) Role of TIF1_ as a modulator of embryonic transcription in the mouse zygote. *Journal of Cell Biology*, 174: 329-38.

Trinkle-Mulcahy, L., Boulon, S., Lam, Y.W., Urcia, R., Boisvert, F.M., Vandermoere, F., Morrice, N.A., Swift, S., Rothbauer, U., Leonhardt, H. and Lamond, A. (2008) Identifying specific protein interaction partners using quantitative mass spectrometry and bead proteomes. *Journal of Cell Biology*, 183: 223-39.

Tse, C., Sera, T., Wolffe, A.P. and Hansen, J.C. (1998) Disruption of higher-order folding by core histone acetylation dramatically enhances transcription of nucleosomal arrays by RNA polymerase III. *Molecular Cellular Biology*, 18: 4629-38.

Tschiersch, B., Hofman, A., Krauss, V., Dorn, R., Korge, G. and Reuter, G. (1994) The protein encoded by the Drosophila position-effect variegation suppressor gene Su(var)3-9 combines domains of antagonistic regulators of homeotic gene complexes. *EMBO Journal*, 13: 3822-31.

Tsukada, Y., Fang, J., Erdjument-Bromage, H., Warren, M.E., Borchers, C.H., Tempst, P. and Zhang, Y. (2006) Histone demethylation by a family of JmjC domain-containing proteins. *Nature*, 439: 811-6.

Tullis, K.M., Krebs, C.J., Leung, J.Y. and Robins, D.M. (2003) The regulator of sex-limitation gene, *rsl*, enforces male-specific liver gene expression by negative regulation. *Endocrinology*, 144: 1854-60.

Ura, K., Hayes, J.J. and Wolffe, A.P. (1995) A positive role for nucleosome mobility in the transcriptional activity of chromatin templates: restriction by linker histones. *EMBO Journal*, 14: 3752-65.

Vakoc, C.R., Letting, D.L., Gheldof, N., Sawado, T., Bender, M.A., Groudine, M., Weiss, M.J., Dekker, J. and Blobel, G.A. (2005) Proximity among distant regulatory elements at the _-globin locus requires GATA-1 and FOG-1. *Molecular Cell*, 17: 453-62.

Van Heyningen, S. and Baxty, B.A. (1997) An ADP-ribosyltransferase from bovine erythrocytes apparently specific for cysteine residues. *Advances in Experimental Medicine and Biology*, 419: 275-8.

Vembar, S.S. and Brodsky, J.L. (2008) One step at a time: endoplasmic reticulum-associated degradation. *Nature Reviews Molecular Cell Biology*, 9: 944-57.

Venter, J.C. *et al* (2001) The sequence of the human genome. *Science*, 291: 1304-51.

Venturini, L., You, J., Stadler, M., Galien, R., Lallemand, V., Koken, M.H., Mattei, M.G., Ganser, A., Chambon, P., Losson, R. and de Thé, H. (1999) TIF1_, a novel member of the transcriptional intermediary factor 1 family. *Oncogene*, 18: 1209-17.

Vertegaal, A.C., Andersen, J.S., Ogg, S.C., Hay, R.T., Mann, M. and Lamond, A.I. (2006) Distinct and overlapping sets of SUMO-1 and SUMO-2 target proteins revealed by quantitative proteomics. *Molecular and Cellular Proteomics*, 5: 2298-310.

Vichi, A., Payne, D.M., Pacheco-Rodriguez, G., Moss, J. and Vaughan, M. (2005) E3 ubiquitin ligase activity of the trifunctional ARD1 (ADP-ribosylation factor domain protein 1). *Proceedings of the National Academy of Sciences USA*, 102: 1945-50.

Visa, N., Puvion-Dutilleul, F., Harper, F., Bachellerie, J.P. and Puvion, E. (1993) Intranuclear distribution of poly(A) RNA determined by electron microscope *in situ* hybridization. *Experimental Cell Research*, 208: 19-34.

Vissing, H., Meyer, W.K., Aagaard, L., Tommerup, N. and Thiesen, H.J. (1995) Repression of transcriptional activity by heterologous KRAB domains present in zinc finger proteins. *FEBS Letters*, 369: 153-7.

Vogel, M.J., Guelen, L., de Wit, E., Peric-Hupkes, D., Lodén, M., Talhout, W., Feenstra, M., Abbas, B., Classen, A.K. and van Steensel, B. (2006) Human heterochromatin proteins form large domains containing KRAB-ZNF genes. *Genome Research*, 16: 1493-504.

Volpi, E.V. *et al* (2000) Large-scale chromatin organization of the major histocompatibility complex and other regions of human chromosome 6 and its response to interferon in interphase nuclei. *Journal of Cell Science*, 113: 1565-76.

Voncken, J.W., Roelen, B.A., Roefs, M., de Vries, S., Verhoeven, E., Marino, S., Deschamps, J. and van Lohuizen, M. (2003) Rnf2 (Ring1b) deficiency causes gastrulation arrest and cell cycle inhibition. *Proceedings of the National Academy of Sciences USA*, 100:2468-73.

Wang, Z.G., Delva, L., Gaboli, M., Rivi, R., Giorgio, M., Cordon-Cardo, C., Grosveld, F. and Pandolfi, P.P. (1998) Role of PML in cell growth and the retinoic acid pathway. *Science*, 279: 1547-51.

Wang, H., Wang, L., Erdjument-Bromage, H., Vidal, M., Tempst, P., Jones, R.S. and Zhang, Y. (2004a) Role of histone H2A ubiquitination in Polycomb silencing. *Nature*, 431: 873-8.

Wang, Z.G., Ruggero, D., Ronchetti, S., Zhong, S., Gaboli, M., Rivi, R. and Pandolfi, P.P. (2004b) Pml is essential for multiple apoptotic pathways. *Nature Genetics*, 20: 266-72.

Watanabe, A., Higuchi, M., Fukushi, M., Ohsawa, T., Takahashi, M., Oie, M. and Fujii, M. (2007) A novel KRAB-zinc finger protein interacts with latency-associated nuclear antigen of Kaposi's sarcoma-associated herpesvirus and activates transcription via terminal repeat sequences. *Virus Genes*, 34: 127-36.

Weinmann, A.S., Yan, P.S., Oberley, M.J., Huang, T.H. and Farnham, P.J. (2002) Isolating human transcription factor targets by coupling chromatin immunoprecipitation and CpG island microarray analysis. *Genes and Development*, 16: 235-44.

Welchman, R.L., Gordon, C. and Mayer, R.J. (2005) Ubiquitin and ubiquitin-like proteins as multifunctional signals. *Nature Reviews Molecular Cell Biology*, 6: 599-609.

Whetstine, J.R., Nottke, A., Lan, F., Huarte, M., Smolikov, S., Chen, Z., Spooner, E., Li, E., Zhang, G., Colaiacovo, M. and Shi, Y. (2006) Reversal of histone lysine trimethylation by the JMJD2 family of histone demethylases. *Cell*, 125: 467-81.

White, D.E., Negorev, D., Peng, H., Ivanov, A.V., Maul, G.G. and Rauscher, F.J.3rd (2006) KAP1, a novel substrate for PIKK family members, colocalizes with numerous damage response factors at DNA lesions. *Cancer Research*, 66: 11594-9.

Williams, R.R., Broad, S., Sheer, D. and Ragoussis, J. (2002) Subchromosomal positioning of the epidermal differentiation complex (EDC) in keratinocyte and lymphoblast interphase nuclei. *Experimental Cell Research*, 272: 163-75.

Williams, R.R., Azuara, V., Perry, P., Sauer, S., Dvorkina, M., Jørgensen, H., Roix, J., McQueen, P., Misteli, T., Merkenschlager, M. and Fisher, A.G. (2006) Neural induction promotes large-scale chromatin reorganisation of the Mash1 locus. *Journal of Cell Science*, 119: 132-40.

Witze, E.S., Old, W.M., Resing, K.A. and Ahn, N.G. (2007) Mapping protein post-translational modifications with mass spectrometry. *Nature Methods*, 4: 798-806.

Witzgall, R., O'Leary, E., Gessner, R., Ouellette, A.J. and Bonventre, J.V. (1993) Kid-1, a putative renal transcription factor: regulation during ontogeny and in response to ischemia and toxic injury. *Molecular and Cellular Biology*, 13: 1933-42.

Wolf, D. and Goff, S.P. (2007) TRIM28 mediates primer binding site-targeted silencing of murine leukemia virus in embryonic cells. *Cell*, 131: 46-57.

Wolf, D., Cammas, F., Losson, R. and Goff, S.P. (2008a) Primer binding site-dependent restriction of murine leukaemia virus requires HP1 binding by TRIM28. *Journal of Virology*, 82: 4675-9.

Wolf, D., Hug, K. and Goff, S.P. (2008b) TRIM28 mediates primer binding site-targeted silencing of Lys1,2 tRNA-utilizing retroviruses in embryonic cells. *Proceedings of the National Academy of Sciences USA*, 105: 12521-6.

Wolf, D. and Goff, S.P. (2009) Embryonic stem cells use ZFP809 to silence retroviral DNAs. *Nature*, doi:10.1038/nature07844

Wondrak, G.T., Cervantes-Laurean, D., Jacobson, E.L. and Jacobson, M.K. (2000) Histone carbonylation *in vivo* and *in vitro*. *Biochemical Journal*, 351: 769-77.

Wysocka, J., Myers, M.P., Laherty, C.D., Eisenman, R.N. and Herr, W. (2003) Human Sin3 deacetylase and trithorax-related Set1/Ash2 histone H3-K4 methyltransferase are tethered together selectively by the cell-proliferation factor HCF-1. *Genes and Development*, 17: 896-911.

Wysocka, J., Swigut, T., Xiao, H., Milne, T.A., Kwon, S.Y., Landry, J., Kauer, M., Tackett, A.J., Chait, B.T., Badenhorst, P., Wu, C. and Allis, C.D. (2006) A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling. *Nature*, 442: 86-90.

Xiang, Y., Zhu, Z., Han, G., Ye, X., Xu, B., Peng, Z., Ma, Y., Yu, Y., Lin, H., Chen, A.P. and Chen, C.D. (2007) JARID1B is a histone H3 lysine 4 demethylase up-regulated in prostate cancer. *Proceedings of the National Academy of Sciences USA*, 104: 19226-31.

Xu, Z.X., Timanova-Atanasova, A., Zhao, R.X. and Chang, K.S. (2003) PML colocalizes with and stabilizes the DNA damage response protein TopBP1. *Molecular and Cellular Biology*, 23: 4247-56.

Yamane, K., Toumazou, C., Tsukada, Y., Erdjument-Bromage, H., Tempst, P., Wong, J. and Zhang, Y. (2006) JHDM2A, a JmjC-containing H3K9 demethylase, facilitates transcription activation by androgen receptor. *Cell*, 125: 483-95.

Yamane, K., Tateishi, K., Klose, R.J., Fang, J., Fabrizio, L.A., Erdjument-Bromage, H., Taylor-Papadimitriou, J., Tempst, P. And Zhang, Y. (2007) PLU-1 is an H3K4 demethylase involved in transcriptional repression and breast cancer cell proliferation. *Molecular Cell*, 25: 801-12.

Yan, K.P., Dollé, P., Mark, M., Lerouge, T., Wendling, O., Chambon, P. and Losson, R. (2004) Molecular cloning, genomic structure, and expression analysis of the mouse transcriptional intermediary factor 1 gamma gene. *Gene*, 334: 3-13.

Yeager, T.R., Neumann, A.A., Englezou, A., Huschtscha, L.I., Noble, J.R. and Reddel, R.R. (1999) Telomerase-negative immortalized human cells contain a novel type of promyelocytic leukemia (PML) body. *Cancer Research*, 59: 4175-9.

Yokoyama, A., Wang, Z., Wysocka, J., Sanyal, M., Auferio, D.J., Kitabayashi, I., Herr, W. And Cleary, M.L. (2004) Leukaemia proto-oncoprotein MLL forms a SET1-like histone methyltransferase complex with menin to regulate Hox gene expression. *Molecular and Cellular Biology*, 24: 5639-49.

Yu, W. *et al* (2004) Poly(ADP-ribosyl)ation regulates CTCF-dependent chromatin insulation. *Nature Genetics*, 36: 1105-10.

Yuan, G.C., Liu, Y.J., Dion, M.F., Slack, M.D., Wu, L.F., Altschuler, S.J. and Rando, O.J. (2005) Genome-scale identification of nucleosome positions in *S. cerevisiae*. *Science*, 309: 626-30.

Zegerman, P., Canas, B., Pappin, D. and Kouzarides, T. (2002) Histone H3 lysine 4 methylation disrupts binding of nucleosome remodelling and deacetylase (NuRD) repressor complex. *Journal of Biological Chemistry*, 277: 11621-4.

Zeng, L., Yap, K.L., Ivanov, A.V., Wang, X., Mujtaba, S., Plotnikova, O., Rauscher, F.J.3rd and Zhou, M.M. (2008) Structural insights into human KAP1 PHD finger-bromodomain and its role in gene silencing. *Nature Structural and Molecular Biology*, 15: 626-33.

Zhang, H., Saitoh, H. and Matunis, M.J. (2002) Enzymes of the SUMO modification pathway localize to filaments of the nuclear pore complex. *Molecular and Cellular Biology*, 22: 6498-508.

Zheng, L., Pan, H., Li, S., Flesken-Nikitin, A., Chen, P.L., Boyer, T.G. and Lee, W.H. (2000) Sequence-specific transcriptional corepressor function for BRCA1 through a novel zinc finger protein, ZBRK1. *Molecular Cell*, 6: 757-68.

Zhong, S., Müller, S., Ronchetti, S., Freemont, P.S., Dejean, A. and Pandolfi, P.P. (2000) Role of SUMO-1 modified PML in nuclear body formation. *Blood*, 95: 2748-52.

Zhou, Y. and Grummt, I. (2005) The PHD finger/bromodomain of NoRC interacts with acetylated histone H4K16 and is sufficient for rDNA silencing. *Current Biology*, 15: 1434-8.

Zhu, J., Chen, Z., Lallemand-Breitenbach, V. and de Thé, H. (2002) How acute promyelocytic leukaemia revived arsenic. *Nature Reviews Cancer*, 2: 705-13.

Zhu, B., Zheng, Y., Pham, A.D., Mandal, S.S., Erdjument-Bromage, H., Tempst, P. and Reinberg, D. (2005) Monoubiquitination of human histone H2B: the factors involved and their roles in HOX gene regulation. *Molecular Cell*, 20: 601-11.

Zink, D., Amaral, M.D., Englmann, A., Lang, S., Clarke, L.A., Rudolph, C., Alt, F., Luther, K., Braz, C., Sadoni, N., Rosenecker, J. and Schindelhauer, D. (2004) Transcription-dependent spatial arrangements of CFTR and adjacent genes in human cell nuclei. *Journal of Cell Biology*, 166: 815-25.

Ziv, Y., Bielopolski, D., Galanty, Y., Lukas, C., Taya, Y., Schultz, D.C., Lukas, J., Bekker-Jensen, S., Bartek, J. and Shiloh, Y. (2006) Chromatin relaxation in response to DNA double-strand breaks is modulated by a novel ATM- and KAP-1 dependent pathway. *Nature Cell Biology*, 8: 870-6.

Appendix A: Publication of Work Presented in this Thesis

KRAB zinc-finger proteins localise to novel KAP1-containing foci that are adjacent to PML nuclear bodies

Stephanie Briers*, Catherine Crawford*, Wendy A. Bickmore and Heidi G. Sutherland‡

MRC Human Genetics Unit, Institute of Genetics and Molecular Medicine, University of Edinburgh, Crewe Road, Edinburgh EH4 2XU, UK

*These authors contributed equally to this work

‡Author for correspondence (e-mail: h.sutherland@hgu.mrc.ac.uk)

Accepted 19 November 2008
Journal of Cell Science 122, 937-946 Published by The Company of Biologists 2009
doi:10.1242/jcs.034793

Summary

The KRAB-zinc finger proteins (KRAB-ZFPs) represent a very large, but poorly understood, family of transcriptional regulators in mammals. They are thought to repress transcription via their interaction with KRAB-associated protein 1 (KAP1), which then assembles a complex of chromatin modifiers to lay down histone marks that are associated with inactive chromatin. Studies of KRAB-ZFP/KAP1-mediated gene silencing, using reporter constructs and ectopically expressed proteins, have shown colocalisation of both KAP1 and repressed reporter target genes to domains of constitutive heterochromatin in the nucleus. However, we show here that although KAP1 does indeed become recruited to pericentric heterochromatin during differentiation of mouse embryonic stem (ES) cells, endogenous KRAB-ZFPs do not. Rather, KRAB-ZFPs and KAP1 relocalise to novel nucleoplasmic foci that we have termed KRAB- and KAP1-associated (KAKA) foci. HP1s can also concentrate in these foci and there is a close spatial relationship between KAKA nuclear foci and PML nuclear bodies. Finally, we reveal differential requirements for the recruitment of KAP1 to pericentric heterochromatin and KAKA foci, and suggest that KAKA foci may contain sumoylated KAP1 – the form of the protein that is active in transcriptional repression.

Supplementary material available online at
http://jcs.biologists.org/cgi/content/full/122/7/937/DC1

Key words: Chromatin, Heterochromatin, Histone methylation, HP1, Nuclear organisation, Transcriptional repression

Introduction

Zinc-finger proteins (ZFPs) that contain the Krüppel-associated box (KRAB) domain comprise the largest single family of transcriptional regulators in the mammalian genome (Huntley et al., 2006; Ravasi et al., 2003; Urrutia, 2003). The basis for their rapid evolution and expansion in mammalian lineages (there are ~400 members in the human or mouse genomes), as well as the target genes that they may regulate, remain poorly understood (Krebs et al., 2005; O'Geen et al., 2007). KRAB-ZFPs are thought to repress transcription via interaction with KRAB-associated protein 1 (KAP1)/TIF1β/TRIM28. In turn, KAP1 is thought to function as a co-repressor by assembling a complex with HP1 (Ryan et al., 1999) and the chromatin modifying enzymes: SETDB1 H3-K9 histone methyltransferase (HMTase) (Schultz et al., 2002), NuRD (Schultz et al., 2001) and histone deacetylases (HDACs) (Nielsen et al., 1999).

Because of a paucity of known physiological targets of KRAB-ZFPs, most progress towards understanding their mechanism of action has come from studies using reporter genes. KRAB domains, which are tethered to a transgene, can repress transcription via KAP1. KAP1 recruits SETDB1 and HP1α to a region around the promoter, resulting in stable transgene silencing (Ayyanathan et al., 2003). This suggests that KRAB-ZFPs might repress gene expression by a localised alteration of chromatin structure and H3K9 methylation. However, the silenced reporter transgene was also seen to re-localise to domains of pericentromeric heterochromatin in the nucleus, suggesting that KRAB-KAP1 repression mechanism may also operate at the level of nuclear organisation.

This latter idea is substantiated by reported changes in the subnuclear distribution of KAP1. During the retinoic acid (RA)-induced differentiation of F9 embryonal carcinoma (EC) and embryonic stem (ES) cells, KAP1 relocalises from the nucleoplasm to the pericentromeric heterochromatin (Cammas et al., 2002). It has been suggested that this might also recruit target genes, and, by implication, the KRAB-ZFPs, to these sites (Cammas et al., 2004). Epitope-tagged KRAB-ZFPs have indeed been reported to be concentrated at pericentromeric heterochromatin in some cells (Matsuda et al., 2001); (Payen et al., 1998; Sutherland et al., 2001). However, the subcellular localisation of endogenous KRAB-ZFP proteins has not been extensively studied.

Here, we have investigated the nuclear distribution of gene-trapped KRAB-ZFPs during the RA-induced differentiation of embryonic stem (ES) cells. We show that, upon differentiation, these fusion proteins, which all retain a KRAB domain, but have lost their zinc fingers, are recruited to pericentromeric heterochromatin in a KAP1-dependent manner. However, we show that these gene-trapped KRAB-ZFPs are mislocalised. Endogenous KRAB-ZFPs, Zfp647 and NT2 do not localise to the domains of pericentric heterochromatin, but, rather, they locate at discrete foci in the nuclei of differentiated ES cells, and these overlap with non-pericentromeric foci of KAP1 and with HP1 proteins. We have termed these KRAB and KAP1-associated (KAKA) foci.

We demonstrate that, whereas the pericentromeric localisation of KAP1 is dependent on trimethylation of H3-K9 that is catalysed by Suv39h1/h2, the formation of KAKA foci is not. We also

establish that there is close spatial proximity between KAKA foci and PML-nuclear bodies (PML-NBs). We suggest that KAKA foci represent a novel nuclear domain involved in the post-translational modification of KRAB-ZFPs and KAP1 by sumoylation, and that they may also be the sites of KRAB-ZFP-mediated gene silencing in the nucleus.

Results

Gene-trapped KRAB-ZFPs relocate to pericentric heterochromatin upon differentiation

We have previously described how a gene-trap screen (which uses an ATG- and promoter-less gene-trap vector, and relies on in-frame splicing into an endogenous gene transcript) can be used to identify proteins that reside in different sub-nuclear compartments (Sutherland et al., 2001). One of the genes trapped in a continuation of this screen was *Zfp647* in the cell line ES492. Sequence analysis indicated that the protein encoded by *Zfp647* is a KRAB A+B family member (Shannon et al., 2003), which is separated by a linker region of 96 amino acids from 13 C_2H_2 zinc fingers. Although the majority of gene-traps do integrate into introns, in the ES492 cell line the vector has inserted into the 5′ end of the final ZF-encoding exon (after amino acid 147) of *Zfp647*, and uses a cryptic splice donor. A schematic representation of Zfp647 and the trapping of this locus, is shown in Fig. 1A. Our previously described gene-trapped KRAB-domain containing proteins showed a nuclear diffuse localisation in ES cells. However, these gene-trapped proteins were also found concentrated at domains of constitutive heterochromatin (which is pericentromeric in mouse) in a subset of ES cells, the morphology of which suggested they may be differentiated (Sutherland et al., 2001).

To determine whether gene-trapped KRAB proteins do relocalise in the nucleus upon differentiation, we used immunofluorescence with an antibody (α-β-gal) that recognises the β-galactosidase (β-gal) region of the gene-trapped KRAB proteins to analyse their nuclear distribution before and after differentiation with retinoic acid (RA). Undifferentiated cells were identified by co-immunofluorescence with an antibody that detects stage-specific embryonic antigen-1 (SSEA-1) on the cell surface (Matsui et al., 1992) [data not shown]. In undifferentiated ES492 gene-trapped cells, the Zfp647-β-gal fusion protein was found in fine speckles distributed across the nucleus, but excluded from pericentric heterochromatin, and mostly excluded from nucleoli (Fig. 1B). However, with differentiation, Zfp647-βgal was found to concentrate at pericentric heterochromatin (brightly staining DAPI foci) in a significant proportion of cells (Fig. 1B,D). We observed (data not shown) a similar relocalisation of β-gal fusion proteins to heterochromatin upon RA-induced differentiation for five other ES cell lines with gene-traps of different KRAB-ZFPs (Sutherland et al., 2001).

This is reminiscent of the differentiation-dependent relocalisation of KAP1 to heterochromatin reported in EC and ES cells (Cammas et al., 2002). To determine whether KAP1 and KRAB-β-gal fusion proteins moved to heterochromatin together, we analysed their subnuclear distribution by co-immunofluorescence, both before and after RA-induced differentiation. In undifferentiated ES cells, KAP1 also showed staining in fine speckles distributed across the nucleus, but excluded from pericentric heterochromatin and nucleoli, although

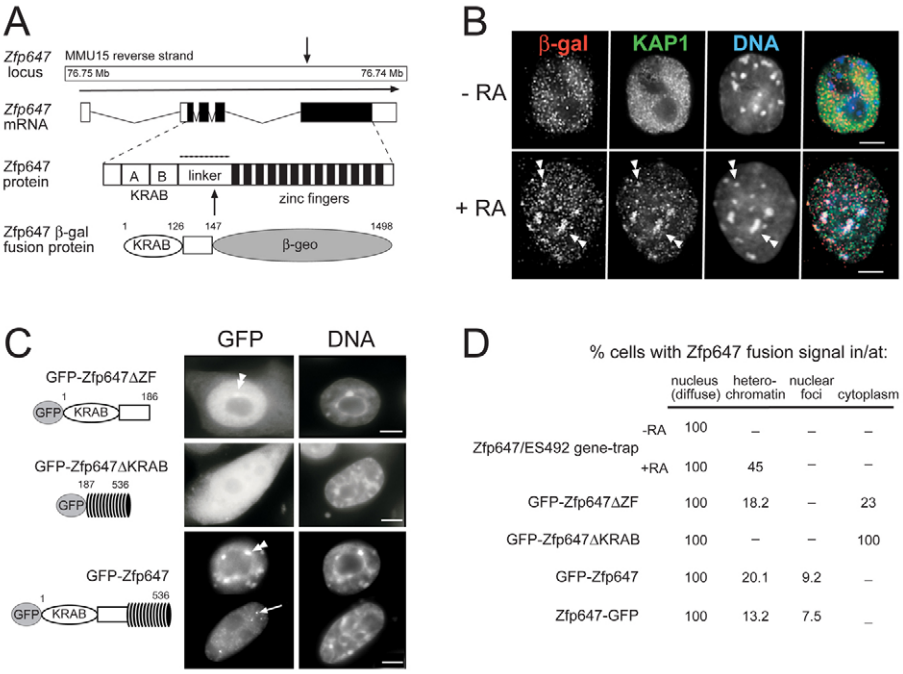


Fig. 1. Gene-trap of Zfp647 and subcellular localisation of Zfp647 fusion proteins with β-gal and GFP. (A) Schematic diagram of gene-trap of *Zfp647* locus in the ES492 cell line. Arrows indicate position of gene-trap integration into the *Zfp647* locus to generate a fusion protein with β-geo (β-gal and neomycin). (B) Deconvolved images from single optical sections of immunofluorescence on ES492 cells before (–RA) and after (+RA) 6 days of differentiation with retinoic acid. Gene-trapped Zfp647 was detected with antibody that detects β-gal (red in merge). Co-staining was with an antibody that recognises KAP1 (green in merge). DNA was counterstained with DAPI (blue in merge). Double arrowheads indicate colocalisation of Zfp647-β-gal fusion protein with KAP1 at pericentric heterochromatin. Scale bars: 5 μm. (C) NIH3T3 cells transiently transfected with: GFP-tagged Zfp647 construct lacking the zinc fingers (GFP-ZfpΔZF); a construct lacking the KRAB domain (GFP-Zfp647ΔKRAB); or full-length Zfp647 (GFP-Zfp647). GFP signal is on the left, DAPI staining is on the right. Double arrowheads show concentrations of GFP-Zfp647 at constitutive heterochromatin (DAPI-bright foci). Arrow indicates a concentration of GFP-Zfp647 at a site that does not correspond with constitutive heterochromatin. Scale bars: 5 μm. (D) Quantification of the various subcellular localisation patterns of β-gal- and GFP-tagged Zfp647.

no colocalisation with Zfp647-β-gal speckles was detected (Fig. 1B, top panel). With differentiation, both KAP1 and KRAB-β-gal fusion proteins localised together at heterochromatin in many cells (Fig. 1B, bottom panel). We never saw KRAB-fusion proteins concentrated at heterochromatin unless KAP1 was present there, suggesting that the relocalisation of KRAB-fusion proteins to pericentric heterochromatin upon differentiation may be dependent on the prior nuclear re-localisation of KAP1 there (Cammas et al., 2002).

Re-localisation of KRAB-ZFPs and KAP1 to heterochromatin after addition of RA could result from differentiation per se, or from changes in the cell cycle as rapidly dividing ES cells differentiate. Therefore, we analysed the localisation of gene-trapped KRAB-ZFPs in BrdU pulse-labelled cells. Before and after differentiaton, we could see KRAB-β-gal fusion proteins concentrated at heterochromatin in both BrdU-positive and negative cells (data not shown). Therefore, KRAB-ZFP movement to heterochromatin is not just a consequence of withdrawal from the cell cycle.

Localisation of GFP-tagged Zfp647

All of our gene-trapped KRAB-ZFPs retain the KAP1-interacting KRAB domain, but lack the zinc fingers (ZFs) of the endogenous protein. As the ZFs probably target the proteins to specific genes, or are responsible for interaction with other proteins, the gene-trapped proteins may mislocalise relative to their wild-type counterparts. We therefore generated constructs of GFP fused with full-length Zfp647, as well as deletion constructs to investigate domains responsible for subcellular localisation.

A construct consisting of GFP fused to Zfp647 lacking the ZFs (GFP-Zfp647ΔZF), thus similar to the gene-trapped protein, could localise to pericentric heterochromatin cells (Fig. 1C, double arrowheads) when transiently expressed in NIH3T3 cells, but also showed some cytoplasmic localisation. The ZFs of Zfp647 alone (GFP-Zfp647ΔKRAB) were distributed diffusely throughout the cell, with no concentration at heterochromatin (Fig. 1C,D). This indicates that amino acids 1 to 147, which comprise the KRAB domain and some of the linker region, are necessary for retention of Zfp647 in the nucleus and its recruitment to domains of heterochromatin. Full-length Zfp647 tagged with GFP at either the N or C terminus (GFP-Zfp647 or Zfp647-GFP) produced a nuclear diffuse staining pattern, but in up to 20% of cells the fusion protein also colocalised with foci of pericentric heterochromatin (Fig. 1C, double arrowhead). This is consistent with data for other transfected epitope-tagged KRAB ZFPs (Matsuda et al., 2001).

Interestingly, in some transfected cells, the GFP-Zfp647 or Zfp647-GFP fusion proteins were also concentrated in smaller foci, which were not at pericentric heterochromatin (arrowed in Fig. 1C and quantified in Fig. 1D).

Generation of antibodies against endogenous Zfp647

The subnuclear localisation of GFP-tagged Zfp647 is generally consistent with that of the gene-trapped KRAB-ZFPs, and with previous studies of KAP1 (Cammas et al., 2002), except that the smaller foci seen with GFP-tagged protein were not apparent in the gene-trapped cells. Therefore, we wished to examine the subcellular localisation of endogenous Zfp647. We raised antibodies specific to amino acids 90-174 in the linker region (broken line above linker of Zfp647 protein diagrammed in Fig. 1A), as this is the region that is most divergent between different KRAB-ZFPs. By western blot, the affinity-purified antibody (α-Zfp647) raised in sheep detected a band close to the expected size of 60 kDa in extracts from wild-type 12.5 days postcoitum (dpc) mouse embryos, and this band was much reduced in intensity in homozygous mutant animals (–/–) generated from the ES492 gene-trap ES cells (Fig. 2A). We believe that the residual band may represent some splicing around the gene-trap integration, rather than crossreaction with any another KRAB ZFPs, as *Zfp647* is not in a cluster of closely related KRAB ZFPs and there is little homology between *Zfp647* and any other KRAB ZFPs across the linker. However, as the antibody raised in sheep was unable to detect the Zfp647 epitope in cells by immunofluorescence (data not shown), we also analysed affinity-purified (α-Zfp647) antibody raised in rabbit. This antibody worked in immunofluorescence and was also able to detect Zfp647 by western blot in extracts from wild-type, but not mutant, animals (Fig. 2B). However, this antibody also recognises a smaller band (~40 kDa), which is non-specific (present in –/–). Subcellular fractionation of NIH3T3 cells reveals that this non-specific band is cytoplasmic, whereas Zfp647 is in the nuclear fraction (Fig. 2B), and therefore it does not interfere with our analysis of Zfp647 staining in the nucleus. Protein loading of the samples for Fig. 2A,B is shown in Fig. 2C, which also confirms the subcellular fractionation of NIH3T3 cells as histones are detected only in the total and nuclear samples, not the cytoplasmic fraction.

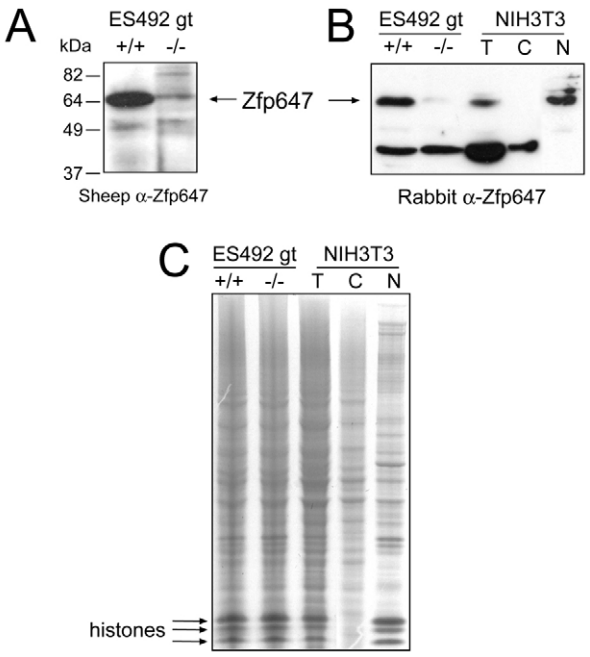


Fig. 2. Detection of endogenous Zfp647 in cell extracts. (A) Western blot with antibody raised in sheep against Zfp647 on extracts from embryos wild-type (+/+) or homozygous mutant (–/–) for the gene-trap into *Zfp647*. (B) Western blot with antibody raised in rabbit against Zfp647 on extracts from embryos wild-type (+/+) or homozygous mutant (–/–) for the gene-trap into *Zfp647*, and in whole cell (T), cytoplasmic (C) and nuclear (N) extracts of mouse NIH3T3 cells. (C) GelCode Blue-stained gel of identical amounts of protein samples as loaded in Fig. 2A,B. The position of the histone proteins is indicated.

Endogenous Zfp647 does not associate with pericentric heterochromatin

We examined the subcellular distribution of endogenous Zfp647 during ES cell differentiation using our rabbit α-Zfp647 antibody. In undifferentiated ES cells (–RA in Fig. 3A), endogenous Zfp647 is dispersed in the nucleoplasm with no obvious overlap with KAP1 staining. In a proportion of ES cells differentiated with RA (+RA),

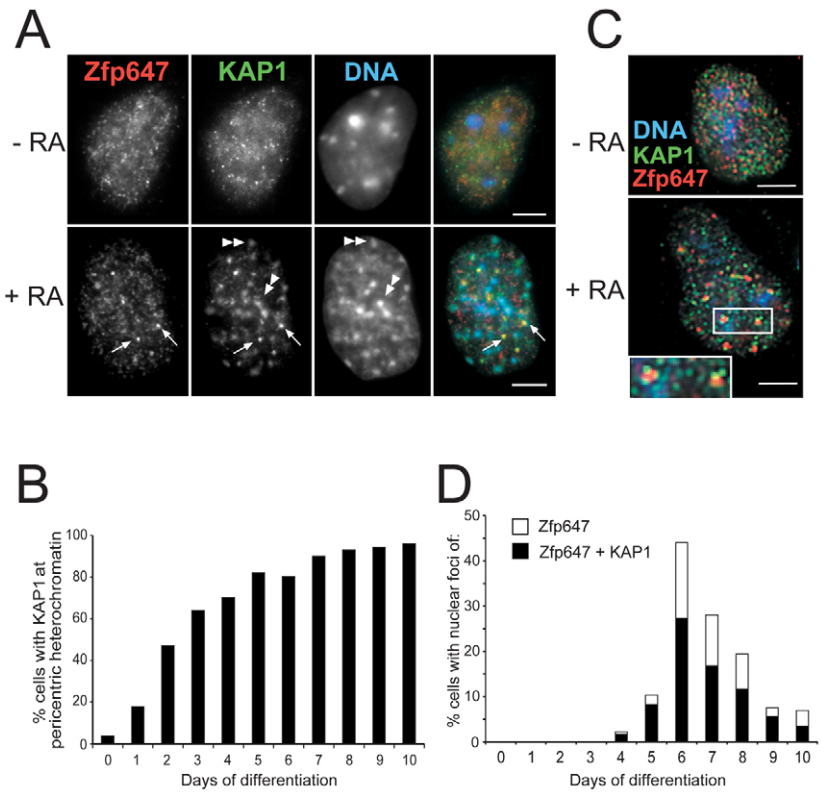


Fig. 3. Detection of endogenous Zfp647 in cells. (A) Immunofluorescence on undifferentiated (–RA) and day 6 differentiated (+RA) OS25 cells with antibodies that recognise Zfp647 (red in merge) and KAP1 (green in merge). DNA was counterstained with DAPI (blue in merge). Double arrowheads indicate KAP1 concentrated at domains of constitutive heterochromatin. Arrows indicate nucleoplasmic foci of Zfp647 and KAP1 that overlap. Scale bars: 5 μm. (B) Percentage of cells with KAP1 localised at the constitutive pericentric heterochromatin with days of differentiation of OS25 ES cells. (C) Deconvolved images from single optical sections of immunofluorescence on undifferentiated (–RA) and day 6 differentiated (+RA) OS25 cells using antibodies that recognise Zfp647 (red) and KAP1 (green). DNA was counterstained with DAPI (blue). Associated nucleoplasmic foci of Zfp647 and KAP1 are shown enlarged in insets. Scale bars: 5 μm. (D) Percentage of cells that show foci of Zfp647 (white bars) and Zfp647 foci associated with KAP1 foci (black bars) with days of differentiation of OS25 ES cells.

Zfp647 relocalised, but unlike the gene-trapped protein, this relocalisation was never to domains of pericentric constitutive heterochromatin (DAPI-bright foci indicated by double arrowheads in Fig. 3A). This may be due to epitope masking of the endogenous protein when it associates with heterochromatin, although we think this is unlikely as α-Zfp647 can detect the same epitope on Zfp647-βgal at heterochromatin in ES492 cells (supplementary material Fig. S1). Instead, in differentiated cells we observed that Zfp647 was recruited to multiple foci (arrowed in Fig. 3A). These are reminiscent of the foci detected with GFP-Zfp647 in transiently transfected cells (Fig. 1C). In the cells in which they appeared, the number of distinct foci of Zfp647 varied from 5 to 40 (average of 14).

Colocalisation of KRAB-ZFPs and KAP1 defines a new nuclear body

Co-staining of ES cells, with antibody that recognises KAP1, showed the relocalisation of KAP1 to sites of constitutive heterochromatin upon differentiation (Fig. 3A, double arrowheads), as has previously been seen by ourselves (Fig. 1B) and others (Cammas et al., 2002). While Cammas et al. (Cammas et al., 2002) found KAP1 to re-localise in a maximum of ~50% cells, and that this was transient in F9 cells, in ES cells we find KAP1 at pericentric heterochromatin increases during differentiation and reaching >90% cells by 7 days of differentiation (Fig. 3B). However, we also detected small nucleoplasmic foci of KAP1, not associated with pericentric heterochromatin, that were coincident with the foci detected by α-Zfp647 (arrowed in Fig. 3A). In deconvolved single image planes, taken at 0.25 μm steps through the *z*-axis of ES cells, both Zfp647 and KAP1 are in fine nuclear speckles in undifferentiated cells that show no obvious overlap (–RA in Fig. 3C). In differentiated cells (+RA), it was apparent that one or more foci of KAP1 may be juxtaposed to, or associated with, each focus of Zfp647. The foci of Zfp647 foci (and their association with KAP1) first appeared around day 4 of ES cell differentiation and are present in a maximum of 44% of cells at day 6 (Fig. 3D). At each time point, in the majority of cells containing Zfp647 foci, they are also associated with KAP1.

We tested whether similar foci exist in other cell types. Non-heterochromatic foci containing both Zfp647 and KAP1 were also seen in the nuclei of primary mouse embryonic fibroblasts (MEFs) from 12.5 dpc embryos (Fig. 4A) and in NIH3T3 cells (data not shown). To determine whether these foci are specific to Zfp647, or whether they might also contain concentrations of other KRAB-ZFPs, we analysed the sub-nuclear localisation of NT2, one of the few other KRAB-ZFPs for which an antibody to the endogenous protein is available (Tanaka et al., 2002). In undifferentiated ES cells, NT2 showed a nucleoplasmic distribution of fine speckles, although a stronger patch of staining was also observed in most cells; we have not identified what this structure is. Upon differentiation with RA, NT2, like Zfp647, relocalised to nucleoplasmic foci that were associated with KAP1 in a subset of cells (arrowed in Fig. 4B) and not to pericentric heterochromatin (Fig. 4B, double arrowheads). Unfortunately, as both antibodies were raised in rabbit we were not able to co-stain with Zfp647 and NT2 to determine whether these two proteins were present in the same foci. We conclude that multiple KRAB-ZFPs can form foci that associate with non-heterochromatic foci of KAP1 in differentiated cells. We propose that these constitute a novel sub-nuclear compartment that we refer to as KRAB and KAP1 associated (KAKA) foci.

KAKA foci also contain concentrations of HP1

KAP1 interacts, through its PxVxL motif, with HP1 proteins (Nielsen et al., 1999; Ryan et al., 1999) and this is an integral part

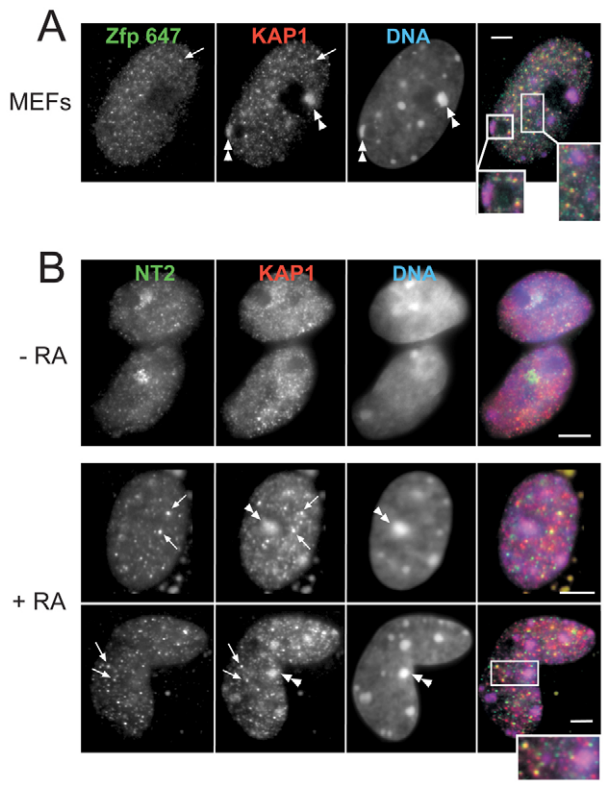


Fig. 4. Colocalisation of NT2 and KAP1 in nucleoplasmic foci. (A) Immunofluorescence on 12.5 dpc primary mouse embryonic fibroblasts (MEFs) with antibodies that recognise Zfp647 (green in merge) and KAP1 (red in merge). DNA was counterstained with DAPI (blue in merge). Double arrowheads indicate KAP1 concentrated at domains of constitutive heterochromatin. Arrows indicate nucleoplasmic foci of Zfp647 and KAP1 that overlap, and others are shown enlarged in inset. Scale bar: 5 μm. (B) Immunofluorescence on undifferentiated (–RA) and day 8 differentiated (+RA) OS25 cells with antibodies that recognise the KRAB-Zfp NT2 (green in merge) and KAP1 (red in merge). DNA was counterstained with DAPI (blue in merge). Double arrowheads indicates KAP1 concentrated at domains of constitutive heterochromatin. Arrows indicate nucleoplasmic foci of KAP1 and NT2, and are shown enlarged in the inset. Scale bars: 5 μm.

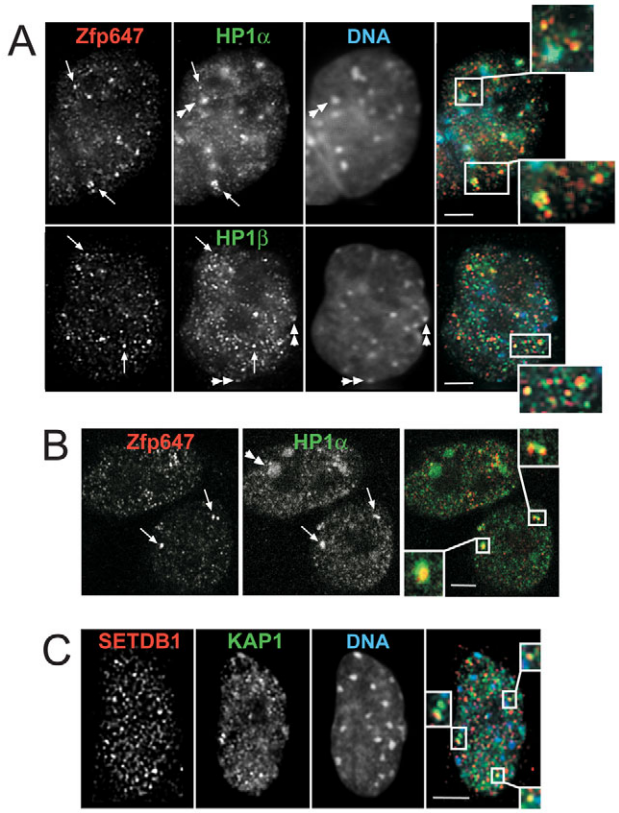


Fig. 5. HP1s concentrate at KAKA foci. (A) Deconvolved images from single optical sections of immunofluorescence on day 6 differentiated OS25 cells with antibodies that recognise Zfp647 (red in merge) and either HP1α or β (green in merge). DNA was counterstained with DAPI (blue in merge). Double arrowheads indicate HP1 concentrated at domains of constitutive heterochromatin. Arrows indicate nucleoplasmic foci of ZFP647 with HP1, and are shown enlarged in insets. (B) Immunofluorescence as in A but in NIH3T3 cells and examined by confocal microscopy. Double arrowhead indicates HP1 concentrated at a domain of constitutive heterochromatin. Arrows indicate nucleoplasmic foci of Zfp647 with HP1, and are shown enlarged in insets. (C) Deconvolved image from single optical section of immunofluorescence on differentiated OS25 cells with antibodies that that recognise SETDB1 (red in merge) and KAP1 (green in merge). DNA was counterstained with DAPI (blue in merge). Nucleoplasmic foci of Zfp647 and SETDB1 that coincide are shown enlarged in insets. Scale bars: 5 μm.

of the mechanism of transcriptional silencing mediated by KRAB-ZFPs/KAP1(Ayyanathan et al., 2003; Sripathy et al., 2006). HP1α and β are concentrated at the domains of heterochromatin where KAP1 is also colocalised (Gilbert et al., 2003); indeed, HP1 has been shown to be required for the localisation of KAP1 to sites of constitutive heterochromatin during differentiation (Cammas et al., 2002). Interestingly, recent FRET data suggest differential interaction between KAP1 and different HP1 isoforms (Cammas et al., 2007).

However, as well as being located at domains of constitutive heterochromatin (Fig. 5A, double arrowheads), we noticed small nucleoplasmic foci of HP1 staining with antibodies that detect HP1α and β. These are associated with the foci demarked by Zfp647 (arrowed in Fig. 5A). Closer examination of the morphology of Zfp647 foci with respect to HP1α and β shows that, similar to KAP1, they are often juxtaposed and also that multiple foci may be linked (Fig. 5A,B, insets). Fig. 5A shows deconvolved single optical sections of differentiated ES cells, but the association of Zfp647 and HP1α in punctate foci was also confirmed in NIH3T3 cells by confocal microscopy (Fig. 5B). We also co-stained Zfp647 with HP1γ in differentiated ES cells, but as HP1γ stained the euchromatic compartment so strongly, we were unable to detect whether it is specifically in KAKA foci (data not shown).

Another protein that can interact with KAP1 and contribute to KRAB-ZFP-mediated transcriptional repression is the H3K9 histone methyltransferase (HMTase) SETDB1 (Ayyanathan et al., 2003; Schultz et al., 2002; Sripathy et al., 2006). By immunofluorescence, we observed that, although in some cells SETDB1 co-localises with KAP1 at pericentric heterochromatin (~40% of cells in day 10 differentiated ES cells, data not shown), most of the protein appears to be located in foci in the nucleoplasm (Fig. 5C). Occasionally, we could detect SETDB1 foci that overlapped with, or were associated with, the KAKA foci detected by immunostaining for KAP1 (Fig. 5C, inset). However, this was not observed with all KAKA foci, suggesting that SETDB1 may not always be present, or that the resident concentration of SETDB1 is not high. Co-

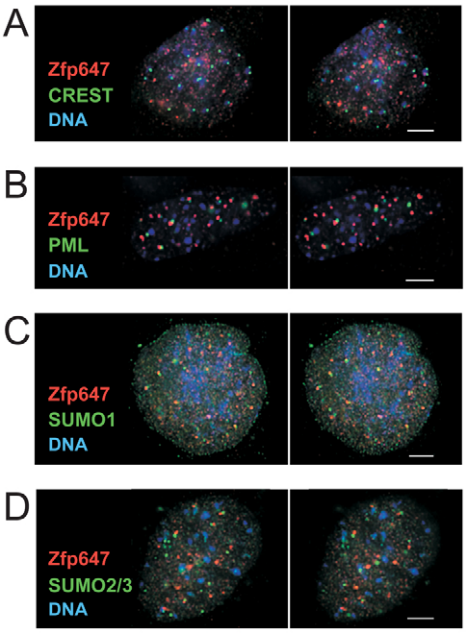


Fig. 6. Association of KAKA foci with PML-NBs. Deconvolved images from two optical sections of day 6 differentiated OS25 cells, separated by 0.5 μm in the *z*-axis, after immunofluorescence with antibody that recognises Zfp647 (red) and, in green, either CREST anti-serum (A), antibody that recognises PML (B), SUMO1 (C) or SUMO2/3 (D). DNA was counterstained with DAPI (blue). Scale bars: 5 μm.

staining with antibodies that detect SETDB1 together with Zfp647 was not possible as both antibodies were raised in rabbit. We suggest that KAKA foci contain complexes of KAP1 with HP1s associated with KRAB-ZFPs, and that SETDB1 may be present.

KAKA foci are often adjacent to PML nuclear bodies

Our data suggest that KRAB-ZFPs and the KAP1/HP1 repression machinery colocalise together at multiple nucleoplasmic foci that we have termed KAKA foci. To determine whether these foci represent a novel nuclear sub-compartment, we analysed their spatial relationship to other nuclear bodies that have a superficial resemblance in size and number.

First, co-staining with CREST antisera, which recognise centromere components, revealed that KAKA foci are distinct from centromeres/kinetochores (Fig. 6A). As KAP1 has been found to colocalise with numerous damage response factors at DNA legions (White et al., 2006), we tested whether foci demarcated by Zfp647 were in fact damage foci using an antibody against γ-H2AX. Although we could not co-stain γ-H2AX with Zfp647 because both antibodies were raised in rabbit, in parallel experiments the foci detected with each antibody were different in appearance and number. Moreover, when we induced DNA damage with UV-C, whereas the number of γ-H2AX foci increased as expected, no change in Zfp647 foci was observed (data not shown).

HP1s, together with MacroH2A, have been found to accumulate in senescence-associated heterochromatin foci (SAHF) (Zhang et al., 2005), but no SAHF bodies were detected with an antibody that recognises MacroH2A during our ES cell differentiation (data not shown).

KAP1 is a member of the TRIM/RBCC (tripartite motif/Ring finger, B-box, coiled-coil) family of proteins, many of which have been described as localised to discrete cellular subcompartments (Meroni and ez-Roux, 2005; Reymond et al., 2001). A prominent such example of an RBCC protein is PML – the principal component and the organiser of PML nuclear bodies (PML-NBs). Co-staining of differentiated ES cells with α-ZFP-647, and with antibody that recognises PML and decorates PML-NBs, revealed that KAKA foci are not coincident with PML-NBs. However, KAKA foci show a spatial relationship with PML-NBs in the majority of the cells in which they are present (arrowed in Fig. 6B). A proportion of Zfp647-containing KAKA foci are in closely apposed pairs with PML-NBs in ~77% of cells in which KAKA foci were detected (varying between 74 and 81% on days 4 to 10 of OS25 ES cell differentiation).

PML was one of the first substrates identified that is subject to modification by the small ubiquitin-like modifier (SUMO) (Sternsdorf et al., 1997) and both SUMOs, and the machinery involved in both the addition and removal of SUMO from substrates, are found concentrated in PML-NBs (Boddy et al., 1996) (http://npd.hgu.mrc.ac.uk/index.html). Indeed, we observed that KAKA foci were also indeed juxtaposed to PML-NBs detected by antibodies that recognise either SUMO1 or SUMO2/3 (Fig. 6C,D).

Targeting of KAP1 to pericentric heterochromatin, but not to KAKA foci, is dependent on Suv39h HMTases

To understand the factors that contribute to the differential targeting of KAP1, HP1 and KRAB-ZFPs to pericentric heterochromatin or to KAKA foci, we took advantage of mouse ES cells that lack the factors responsible for epigenetic marks characteristic of heterochromatin. The highest concentration of methylated CpGs in mouse cells is at pericentric heterochromatin, and the DNA methyltransferases (DNMTs) Dnmt3a and Dnmt3b also concentrate there (Bachman et al., 2001). In the absence of these DNMTs, and with the consequent loss of DNA methylation, there is an altered nuclear organisation and histone modification profile at pericentric heterochromatin (Gilbert et al., 2007). Transgene studies have also implicated DNA methylation and recruitment to domains of pericentric heterochromatin in the KRAB-mediated repression mechanism (Ayyanathan et al., 2003). However, in late passage *Dnmt3a/b*$^{-/-}$ cells, which we have previously shown have no detectable remaining CpG methylation (Gilbert et al., 2007), we found that KAP1 was still concentrated at heterochromatin (data not shown). Therefore, DNA methylation at pericentromeric heterochromatin is not needed to sequester KAP1 there.

Another chromatin mark characteristic of pericentric heterochromatin is tri-methylated H3-K9 (H3K9me3) laid down by the Suv39h family of HMTases. In mouse cells null for both *Suv39h1* and *Suv39h2* (*Suv39h dn*), the absence of H3K9me3 leads to a delocalisation of HP1α and HP1β from pericentromeric heterochromatin (Lehnertz et al., 2003) (Fig. 7A). We also found that KAP1 localisation to pericentric heterochromatin was lost in *Suv39h dn* cells (Fig. 7B), indicating that KAP1 concentration at these nuclear domains may be driven through its interactions with HP1s. However, discrete KAKA foci containing KAP1, HP1 and ZFP647 were still present in the nuclei of *Suv39h dn* cells (Fig. 7). This suggests that formation of KAKA foci is independent of the activity of the Suvar39h HMTases and that the two nuclear sites of accumulation of KAP1 are fundamentally different from one another.

Sumoylation of KAP1 corresponds to the appearance of KAKA foci

We also observed that both SUMO1 and SUMO2/3 could be found concentrated at pericentric heterochromatin in some differentiated

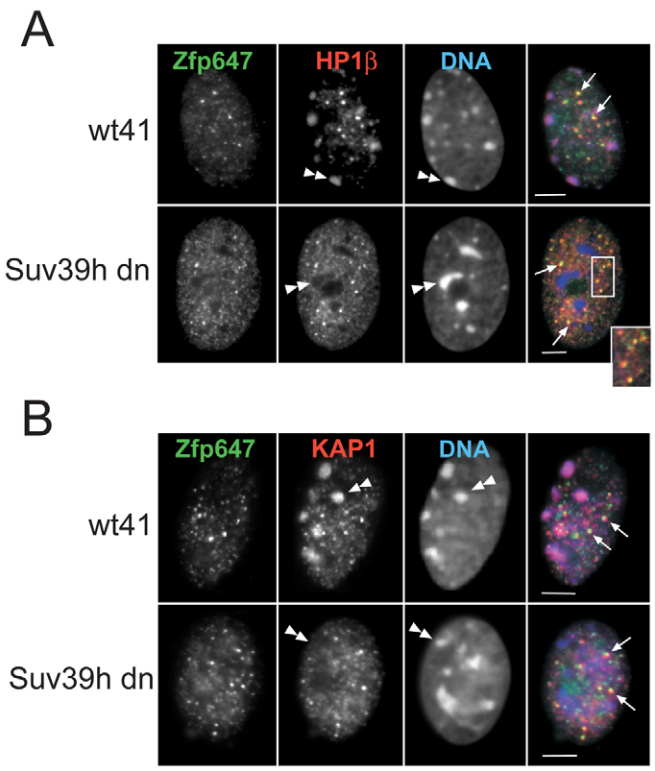


Fig. 7. KRAB-ZFP, HP1 and KAP1 localisation in Suvar39h knockout cells. (A) Immunofluorescence on day 8 differentiated ES cells null for both Suv39h1 and h2 (Suv39h dn), and on the wild-type parental controls for these cells (wt41) with antibodies that detect Zfp647 (green in merge) and HP1β (red in merge). DNA was counterstained with DAPI (blue in merge). Arrows indicate nucleoplasmic foci of Zfp647 and HP1β, and some are shown enlarged in inset. Double arrowheads indicate HP1β concentrated at domains of constitutive heterochromatin in wild-type cells, but delocalised in Suv39h dn cells. Scale bars: 5 µm. (B) As in A but with antibodies that recognise Zfp647 (green in merge) and KAP1 (red in merge). Arrows indicate nucleoplasmic foci of Zfp647 and KAP1. Double arrowheads indicate KAP1 concentrated at domains of constitutive heterochromatin in wild-type cells, but delocalised in Suv39h dn cells. Scale bars: 5 µm.

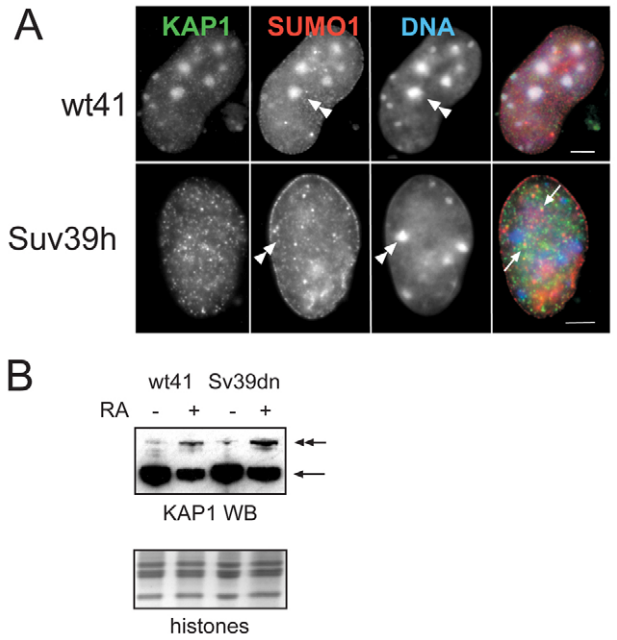


Fig. 8. KAP1 and sumoylation. (A) Immunofluorescence on day 8 differentiated ES cells null for both Suv39h1 and h2 (Suv39h), and on the wild-type parental controls for these cells (wt41) with antibodies that detect KAP1 (green in merge) and SUMO1 (red in merge). Double arrowhead indicates concentration of SUMO1 with domains of constitutive heterochromatin in wild-type cells, but delocalised in Suv39h dn cells. Arrows indicate SUMO1 in PML-NBs that are juxtaposed to KAKA foci. Scale bars: 5 µm. (B) Western blot with antibody that detects KAP1 in extracts prepared from undifferentiated (–RA) and day 10 differentiated (+RA) ES cells null for both Suv39h1 and h2 (Sv39dn), and on the wild-type parental controls for these cells (wt41). Arrow indicates the unmodified form of KAP1 and double arrow indicates sumoylated KAP1. The panel below shows the GelCode Blue-stained histones as a loading control.

ES cells (SUMO1 shown in Fig. 8A). As KAP1 itself can be sumoylated (Ivanov et al., 2007; Lee et al., 2007; Li et al., 2007; Mascle et al., 2007), we thought it possible that the SUMO we observe at pericentric heterochromatin is actually modified KAP1. However, at day 10 of differentiation, the percentage of cells in which SUMO1 and SUMO2/3 is at pericentric heterochromatin (19% and 46%, respectively) is far below that in which KAP1 is at these sites (96%). Interestingly, in *Suv39h dn* cells, pericentric localisation was no longer seen for either SUMO (SUMO1 shown in Fig. 8A), suggesting the localisation of a major sumoylated target or of SUMO itself to this site is dependent on this HMTase or H3K9me3. By contrast, the juxtaposition of KAKA foci and SUMO-containing PML NBs persists in the absence of the Suv39h HMTases (arrowed in Fig. 8A).

The appearance of KAKA foci during ES cell differentiation, and their proximity to SUMO-containing PML-NBs, prompted us to address whether sumoylation of KAP1 is also regulated during differentiation. By western blot with a KAP1 antibody, we analysed total cell extracts, prepared in the presence of *N*-ethylmaleimide (NEM) to preserve SUMO modification, from both undifferentiated (–RA) and day 10 differentiated (+RA) wild-type ES cells (wt41). We found that the major unmodified form of KAP1 (Fig. 8B, single arrow) was downregulated during ES cell differentiation, as previously described for F9 cells (Cammas et al., 2004). However, interestingly, a higher molecular weight form, corresponding to SUMOylated KAP1 (double arrow), was increased with differentiation. This was also seen in extracts from *Suv39 dn* ES cells. This suggests that the sumoylated form of KAP1, which is the form active in transcriptional repression (Ivanov et al., 2007; Lee et al., 2007; Li et al., 2007; Mascle et al., 2007), is resident in the euchromatic compartment, perhaps in KAKA foci, and not at the constitutive heterochromatin, as KAP1 is delocalised from pericentric heterochromatin in *Suv39 dn* ES cells (Fig. 7B).

Discussion

KRAB ZFPs have a dynamic nuclear localisation

The initial studies of transgenes silenced by the KRAB-KAP1 repression mechanism (Ayyanathan et al., 2003), the KAP1 localisation studies during ES and EC cell differentiation (Cammas et al., 2002; Cammas et al., 2004), and the subnuclear localisation of epitope-tagged KRAB-ZFPs (Matsuda et al., 2001; Payen et al., 1998; Sutherland et al., 2001) (Fig. 1) had all pointed towards a link between the KAP1/KRAB-ZFP repression pathway and domains of pericentric heterochromatin in the nucleus. In this paper, we present the first detailed study of the subnuclear distribution of endogenous KRAB-ZFPs. We show that these proteins do have a

dynamic nuclear localisation during cell differentiation, but that they are not seen concentrated at the visible domains of constitutive heterochromatin in differentiated mouse cells, even in cells where KAP1 is concentrated there (Figs 3 and 4).

What is the explanation for the differential localisation of endogenous and ectopically expressed KRAB-ZFPs? Ectopic expression of GFP-tagged Zfp647 deletion constructs show that it is the KRAB domain itself that is required for localisation to heterochromatin (Fig. 2). This is probably due to the specific interaction of the KRAB domain with KAP1 (Friedman et al., 1996) and is supported by previous findings where other tagged KRAB-ZFPs, with mutations in the KRAB domain that are unable to bind KAP1, can no longer localise to centromeric foci (Matsuda et al., 2001). It seems most likely that the ZFs or linker regions of KRAB ZFPs, such as Zfp647, provide the DNA-binding specificity to direct the protein to its correct target loci. We suggest that in the absence of these regions (deletion constructs or gene-trapped proteins), the KRAB-KAP1 interaction dominates and results in a mis-localisation of the KRAB protein at heterochromatin. Ectopic expression of epitope tagged full-length KRAB-ZFPs may saturate specific target binding sites in the genome and so also result in the accumulation of the tagged protein at heterochromatin.

Thus, caution must be exercised in interpreting the mechanism of transcriptional repression when using ectopically expressed KRAB-ZFPs; it remains unclear what the function of KAP1 is, if any, at constitutive heterochromatin. Some KRAB proteins do consist of only a KRAB domain, e.g. MIF1 (Nikulina et al., 2006), and some are alternatively spliced to generate a KRAB-box only protein, e.g. KRAB-O from the *Zfp208* locus (Oh et al., 2005). Therefore, it is possible that localisation at pericentric heterochromatin is relevant to the function of a subset of KRAB box-containing proteins. However, our finding that, in *Suv39h dn* ES cells, KAP1 is completely delocalised from pericentric heterochromatin, presumably as a consequence of HP1 delocalisation (Fig. 7), suggests that the major role of KAP1 is not at pericentric heterochromatin. This is supported by the fact that the phenotype of the mice null for both Suv39h HMTases is quite mild, with all animals developing normally until 12.5 dpc and many mice surviving to adulthood (Peters et al., 2001). A null deletion of KAP1 itself, however, results in complete embryonic lethality just after implantation (Cammas et al., 2000). Ivanov et al. (Ivanov et al., 2007) have reported that a sumoylation-deficient mutant of KAP1 (K6R) still localised to pericentric heterochromatin and suggest that loss of repression conferred by this mutant was not due to a defect in subnuclear targeting. This would also be consistent with our suggestion that pericentric heterochromatin is not the site of KRAB ZFP-mediated gene repression.

KRAB ZFPs characterise a novel nuclear body

In differentiated mouse cells, we found that two endogenous KRAB-ZFPs (Zfp647 and NT2) were concentrated in nucleoplasmic foci that are not associated with pericentric heterochromatin, but are tightly associated with foci of KAP1 (Figs 3 and 4). We have termed these KRAB- and KAP1-associated (KAKA) foci. We consider it likely that the KAP1-containing nucleoplasmic foci reported with the GFP-tagged KRAB ZFP PAROT are also KAKA foci (Fleischer et al., 2006).

KAKA foci also contain the known KAP1-interacting proteins HP1α and HP1β (Nielsen et al., 1999; Ryan et al., 1999). We also show that another component of the KAP1-mediated repression machinery, SETDB1 (Schultz et al., 2002), is present in nucleoplasmic foci that we observe occasionally overlapping with KAKA foci (Fig. 5C). The KAP1-HP1-KRAB-ZFP complexes in KAKA foci may only be transiently interacting with SETDB1, or they may be interacting mainly with other chromatin modifiers at these sites. In this regard, it is noteworthy that only 25% of the KAP1 binding sites identified in the genome of Ntera2 cells are also marked by the presence of H3K9me3 – the histone modification catalysed by SETDB1 (O'Geen et al., 2007).

KAKA foci – twin of PML NBs?

We have shown that KAKA foci are often juxtaposed to PML nuclear bodies (Fig. 6B). This is reminiscent of the relationship between gems, Cajal bodies and cleavage bodies, all structures involved in steps of pre-mRNA processing. These bodies can occur as separate nuclear structures, but can be also be twinned or colocalised in some situations (Li et al., 2006).

Intriguingly, given the spatial juxtaposition of KAKA foci and PML-NBs, both KAP1 and PML are proteins of the RBCC family. Modification of PML by sumoylation is known to be important for the formation of PML-NBs (Zhong et al., 2000) and SUMOs are concentrated in PML-NBs (Bernardi and Pandolfi, 2007) (Fig. 6C,D). As we have found that the appearance of high molecular weight forms of KAP1 during differentiation, consistent with its sumoylation (Fig. 8), is coincident with the cytological appearance of KAKA foci (Figs 3 and 4), we speculate that, similar to PML-NBs, sumoylation of KAP1 is involved in the formation of KAKA foci. Consistent with a link between KAKA foci and sumoylation, KRAB-ZFPs themselves have also been shown to be sumoylated in a proteomics analysis of proteins modified by SUMO1 and SUMO2/3 (Vertegaal et al., 2006).

It has been suggested that PML itself might be a SUMO E3 ligase, modifying both itself and other targets (Bernardi and Pandolfi, 2007; Quimby et al., 2006). Similarly, KAP1 has also been shown to be an intra-molecular SUMO E3 ligase (Ivanov et al., 2007). As it is the SUMOylated form of KAP1 that recruits the downstream effectors such as SETDB1 and NURD, this suggests that KAKA foci might be sites for modification and/or assembly of the components of KRAB-ZFP-mediated transcriptional repression (Ivanov et al., 2007; Li et al., 2007; Mascle et al., 2007).

It remains for bone fide target genes of KRAB-ZFPs, such as Zfp647, to be identified in order to address whether target genes are recruited to KAKA foci for their silencing or whether KAKA foci are storage, processing or assembly sites for the proteins involved in KRAB-ZFP-mediated silencing.

Materials and Methods

Cell culture and differentiation

Undifferentiated E14 wt, gene-trapped, *Dnmt3ab*$^{-/-}$ (Okano et al., 1999), *Suv39h1/h2*$^{-/-}$ (Peters et al., 2001) and OS25 (Billon et al., 2002) ES cells were maintained on 0.1% gelatin-coated dishes in Glasgow's modified Eagle's medium (GMEM) containing 10% foetal calf serum (FCS), non-essential amino acids, 1 mM sodium pyruvate, 0.3 mg/ml L-glutamine, 0.1 mM 2-mercaptoethanol and 1000 U/ml human recombinant LIF. Mouse NIH3T3 fibroblasts were maintained in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% FCS.

100 μg/ml hygromycin B was used to select for undifferentiated (Oct4-expressing) OS25 cells. To induce their differentiation, 5×10^5 OS25 cells were plated in 25 cm^2 flasks (Costar) without LIF or hygromycin B for 1 day. The cells were maintained in medium with 5×10^{-6} M RA for the next 4 days. For the remaining length of differentiation (up to day 10), 2.5 μM gancyclovir was also added to RA-containing medium to select against undifferentiated (Oct4-expressing) cells (Billon et al., 2002). Media was changed every second day and cells were replated or seeded onto 4×10 cm^2 gelatinised slides as necessary. All other cell lines maintained and differentiated as above with RA, but with no selection.

DNA sequence analysis of ES492/Zfp647

5′ RACE was used to obtain ~800 bp of sequence information for the ES492 gene-trap as previously described (Sutherland et al., 2001). Mouse genomic and EST databases were searched for sequence matches using the BLAST algorithm (http://www.ncbi.nih.gov/BLAST/ or http://www.ensembl.org/). This revealed 94% nucleotide identity with hypothetical mouse mRNA (NM_172817), which encodes Zfp647 (NP_766405), the 535 amino acid KRAB-ZFP.

We subcloned and sequenced the full-length Zfp647 cDNA from EST BI656339 (IMAGE clone 5326813). Domains of the trapped proteins were examined using InterProScan (http://www.ebi.ac.uk/InterProScan/) and the Simple Modular Architecture Research Tool (SMART) (http://smart.embl-heidelberg.de/). The KRAB box is of the A+B family (Shannon et al., 2003), and a linker region of 96 amino acids separates the KRAB domain from the 13 C2H2 zinc fingers (ZFs). The gene trap is inserted into the final ZF-encoding exon (after amino acid 147) and uses a cryptic splice donor (http://npd.hgu.mrc.ac.uk/search.php?action=builddetails&geneid=1NP01271).

Cell transfection

Full-length *Zfp647* cDNA, and versions lacking the ZFs, or the KRAB box, were cloned from PCR products in frame into pEGFP-C1 or -N1 vectors (Clontech). Zfp647 lacking the ZFs (GFP-Zfp647ΔZF) consisted of GFP fused to amino acids 1 to 186 of Zfp647. Zfp647 without the KRAB box consisted of GFP fused to amino acids 187 to 535 (Fig. 1C). NIH3T3 cells were transfected with Lipofectamine 2000 (Invitrogen) according to the manufacturer's instructions. Cells were incubated for 24 hours, fixed and visualised for GFP expression.

Bacterial fusion proteins and antibody production

Products encoding amino acids 90-174 of Zfp647, covering most of the linker region, were PCR amplified from IMAGE clone 5326813 using the primers 5′ACTG-AATTCGGAAGTGACCACTCAGAATGC3′ and 5′TGAGCGGCCGCATAGGG-TCTCTCAACAGTGGG3′. The PCR product was subcloned into pGex-4T-1 (Amersham Biosciences), sequenced and transformed into *E. coli* BL21-CodonPlus (DE3)RP cells (Stratagene). Bacterial GST-ZFP647 linker fusion protein was induced with IPTG and purified on glutathione-agarose (Sigma) before injection into rabbits and sheep (Diagnostic Scotland). The same PCR product was cloned into pET32a (Novagen); the His-tagged fusion protein was purified on a Ni-agarose column (HIS-Select Nickel Affinity Gel, Sigma) and an affinity column was made by binding it to CNBr-activated sepharose 4B (Amersham Biosciences). Antibodies recognising Zfp647 (α-Zfp647) were affinity purified against the Zfp647 linker affinity column and specificity was confirmed by western blotting against His-tagged Zfp647 protein expressed in *E. coli* (data not shown).

Subcellular fractionation and western blotting

Protein extracts from 12.5 dpc embryos were prepared by homogenisation of dissected limbs (where Zfp647 is expressed) in 2×SDS loading buffer and boiling samples before loading. Whole-cell protein extracts were made by overlaying cells with PBS, adding an equal volume of 2×SDS protein loading buffer and boiling before loading. NEM was added to the loading at a final concentration of 20 mM in the protein samples in which it was used. For nuclear extracts, cells were washed in PBS and resuspended on ice in nuclei extraction Buffer A [NBA, 5.5% sucrose (w/v in dH_2O), 10 mM Tris-HCl (pH 8), 85 mM KCl, 0.5 mM spermidine, 250 μM PMSF (phenyl methyl sulfonyl fluoride), 0.2 mM EDTA]. An equal volume of NBA + 0.1% NP-40 [v/v] was then added prior to pelleting at 500 ***g*** for 3 minutes at 4°C. Nuclei were washed with NBA and their concentration determined from the A_{260}. An equal volume of 2×SDS protein loading buffer was added to the sample and boiled before loading on SDS-PAGE gels. Total cell, cytoplasmic and nuclear NIH3T3 fractions were prepared as previously described (Sutherland et al., 2004).

For western blotting, protein extracts were fractionated by 10-12% SDS-PAGE and transferred to a nylon membrane by wet blotting (GENIE blotter, Idea Scientific). The membranes were incubated with primary antibodies and detected by horseradish peroxidase (HRP)-conjugated donkey anti-rabbit or anti-mouse whole molecule IgG (Sigma, 1:10,000) and chemiluminescence (SuperSignal 1:2000). Primary antibody dilutions were as follows: 1:1000 sheep α-647 (3.2 mg/ml), 1:1500 rabbit α-647 (4.75 mg/ml) and 1:2500 rabbit α-KAP1 (Bethyl Laboratories, A300-274A). For loading controls, protein extracts were stained with GelCode Blue stain reagent (ThermoScientific) after fractionation. For histone loading controls, protein extracts were fractionated by 17% SDS-PAGE before GelCode Blue staining.

Immunofluorescence

Cells were grown as monolayers on slides and fixed for 20 minutes in 3% paraformaldehyde (pFa)/PBS containing 1.5 mM $MgCl_2$ and 1 mM $CaCl_2$. Fix was quenched in 50 mM NH_4Cl/PBS for 10 minutes and cells were permeabilised in 0.25% Triton X-100/ PBS for 12 minutes. Cells were incubated overnight with primary antibodies: 1:150 dilution of rabbit α-647 (this paper); undiluted mAb supernatant recognising KAP1 (gift of F. J. Rauscher III and D. Schultz) or a 1:1200 dilution of a mAb against KAP1 (gift of P. Chambon); 1:500 dilutions of mouse mAb α-HP1α, β and γ (Chemicon); 1:200 dilution mouse mAb α-SSEA-1 (DSHB); 1:2000 dilution of rabbit α-β-gal (Europa) or a mAb against β-gal (Promega); 1:1000 dilution of mouse mAb α-PML (Chemicon); 1:300 dilution of CREST sera (Gilchrist et al., 2004); 1:150 dilution of sheep α-SUMO1 and 1:50 dilution of α-SUMO2/3 (gifts of Ron Hay); 1:200 dilution of MacroH2A1 (Upstate); and 1:500 dilution of α-phospho-H2AX [Ser139] (Upstate). After washing off non-specifically bound antibody, the slides were then incubated for 1 hour with 1:200 dilutions of FITC- or TexasRed-labelled secondary antibodies (Jackson and Vector Laboratories) or 1:1000 dilutions of Alexa Fluor 488 or Alexa Fluor 594 secondary antibodies (Invitrogen).

For treatment with UV, cells were irradiated with 200 J/m^2 of UV-C at 254 nm (UV Stratalinker 1800, Stratagene), returned to media and incubated for 30 minutes before fixation and immunofluorescence as described above. For immunofluorescence combined with detection of bromodeoxyuridine (BrdU) incorporation, 0.01 M BrdU (Roche) was added to cells in culture 30 minutes before fixation in 3% pFa for 20 minutes. Cells were incubated with primary antibody and subsequently fixed in 10% formalin (v/v)/PBS for 10 minutes and permeabilised with 0.1% Triton X-100/PBS for 12 minutes. DNA was denatured with a 30-minute 2 M HCl treatment to allow for α-BrdU antibody access. Cells were washed in PBS, blocked for 10 minutes in 5% BSA (w/v) before incubating with 1:100 dilution of a rabbit antibody recognising BrdU (Harlan SeraLab) for 1 hour, followed by incubation with secondary antibodies.

Slides were counterstained with 0.5 μg/ml DAPI in Vectashield, and examined on a Zeiss Axioplan epifluorescence microscope equipped with a triple band-pass filter (Chroma #83000) and imaged with cooled CCD camera using IPLAB software v. 3.6 (Scanlytics, USA). For optical sectioning, the microscope objective was fitted with a Pifoc motor to allow optical sectioning in the *z*-axis and the images were subject to deconvolution to remove out of focus blur using Hazebuster deconvolution software. A Zeiss LSM510 laser-scanning confocal microscope was also used for examining some slides.

S.B. and C.C. were funded by PhD studentships from the James S. McDonnell Foundation and the UK Medical Research Council, respectively. W.A.B. is a Centennial fellow of the James S. McDonnell Foundation. H.G.S. was part funded by the AICR. This work was supported by the Medical Research Council, UK and in part by the EU FP6 Network of Excellence Epigenome (LSHG-CT-2004-503433). We thank Phillipe Gautier for bioinformatics assistance. We thank Frank Rauscher III (Wistar Institute, Philadelphia) and D. Schultz (Case Western Reserve University, Cleveland) for the KAP1 mAb; Pierre Chambon (IGBMC, University of Louis Pasteur, France) for the mouse TIF1β mAb; Yoshihiko Yamada (Kyushu University, Japan) for the NT2 antibody; and Ron Hay (Wellcome Trust Biocentre, University of Dundee) for the SUMO1 and SUMO2/3 antibodies. En Li (Novartis Institutes for BioMedical Research) and Thomas Jenuwein (IMP, Vienna) provided Dnmt3ab and Suv39h double knockout ES cells, respectively. Deposited in PMC for release after 6 months.

References

Ayyanathan, K., Lechner, M. S., Bell, P., Maul, G. G., Schultz, D. C., Yamada, Y., Tanaka, K., Torigoe, K. and Rauscher, F. J., 3rd (2003). Regulated recruitment of HP1 to a euchromatic gene induces mitotically heritable, epigenetic gene silencing: a mammalian cell culture model of gene variegation. *Genes Dev.* **17**, 1855-1869.

Bachman, K. E., Rountree, M. R. and Baylin, S. B. (2001). Dnmt3a and Dnmt3b are transcriptional repressors that exhibit unique localization properties to heterochromatin. *J. Biol. Chem.* **276**, 32282-32287.

Bernardi, R. and Pandolfi, P. P. (2007). Structure, dynamics and functions of promyelocytic leukaemia nuclear bodies. *Nat. Rev. Mol. Cell. Biol.* **8**, 1006-1016.

Billon, N., Jolicoeur, C., Ying, Q. L., Smith, A. and Raff, M. (2002). Normal timing of oligodendrocyte development from genetically engineered, lineage-selectable mouse ES cells. *J. Cell Sci.* **115**, 3657-3665.

Boddy, M. N., Howe, K., Etkin, L. D., Solomon, E. and Freemont, P. S. (1996). PIC 1, a novel ubiquitin-like protein which interacts with the PML component of a multiprotein complex that is disrupted in acute promyelocytic leukaemia. *Oncogene* **13**, 971-982.

Cammas, F., Mark, M., Dolle, P., Dierich, A., Chambon, P. and Losson, R. (2000). Mice lacking the transcriptional corepressor TIF1beta are defective in early postimplantation development. *Development* **127**, 2955-2963.

Cammas, F., Oulad-Abdelghani, M., Vonesch, J. L., Huss-Garcia, Y., Chambon, P. and Losson, R. (2002). Cell differentiation induces TIF1beta association with centromeric heterochromatin via an HP1 interaction. *J. Cell Sci.* **115**, 3439-3448.

Cammas, F., Herzog, M., Lerouge, T., Chambon, P. and Losson, R. (2004). Association of the transcriptional corepressor TIF1beta with heterochromatin protein 1 (HP1): an essential role for progression through differentiation. *Genes Dev.* **18**, 2147-2160.

Cammas, F., Janoshazi, A., Lerouge, T. and Losson, R. (2007). Dynamic and selective interactions of the transcriptional corepressor TIF1 beta with the heterochromatin protein HP1 isotypes during cell differentiation. *Differentiation* **75**, 627-637.

Fleischer, S., Wiemann, S., Will, H. and Hofmann, T. G. (2006). PML-associated repressor of transcription (PAROT), a novel KRAB-zinc finger repressor, is regulated through association with PML nuclear bodies. *Exp. Cell Res.* **312**, 901-912.

Friedman, J. R., Fredericks, W. J., Jensen, D. E., Speicher, D. W., Huang, X. P., Neilson, E. G. and Rauscher, F. J., 3rd (1996). KAP-1, a novel corepressor for the highly conserved KRAB repression domain. *Genes Dev.* **10**, 2067-2078.

Gilbert, N., Boyle, S., Sutherland, H., de Las, H. J., Allan, J., Jenuwein, T. and Bickmore, W. A. (2003). Formation of facultative heterochromatin in the absence of HP1. *EMBO J.* **22**, 5540-5550.

Gilbert, N., Thomson, I., Boyle, S., Allan, J., Ramsahoye, B. and Bickmore, W. A. (2007). DNA methylation affects nuclear organization, histone modifications, and linker histone binding but not chromatin compaction. *J. Cell Biol.* **177**, 401-411.

Gilchrist, S., Gilbert, N., Perry, P. and Bickmore, W. A. (2004). Nuclear organization of centromeric domains is not perturbed by inhibition of histone deacetylases. *Chromosome Res.* **12**, 505-516.

Huntley, S., Baggott, D. M., Hamilton, A. T., Tran-Gyamfi, M., Yang, S., Kim, J., Gordon, L., Branscomb, E. and Stubbs, L. (2006). A comprehensive catalog of human KRAB-associated zinc finger genes: insights into the evolutionary history of a large family of transcriptional repressors. *Genome Res.* **16**, 669-677.

Ivanov, A. V., Peng, H., Yurchenko, V., Yap, K. L., Negorev, D. G., Schultz, D. C., Psulkowski, E., Fredericks, W. J., White, D. E., Maul, G. G. et al. (2007). PHD domain-mediated E3 ligase activity directs intramolecular sumoylation of an adjacent bromodomain required for gene silencing. *Mol. Cell* **28**, 823-837.

Krebs, C. J., Larkins, L. K., Khan, S. M. and Robins, D. M. (2005). Expansion and diversification of KRAB zinc-finger genes within a cluster including *Regulator of sex-limitation 1 and 2*. *Genomics* **85**, 752-761.

Lee, Y. K., Thomas, S. N., Yang, A. J. and Ann, D. K. (2007). Doxorubicin down-regulates Kruppel-associated box domain-associated protein 1 sumoylation that relieves its transcription repression on p21WAF1/CIP1 in breast cancer MCF-7 cells. *J. Biol. Chem.* **282**, 1595-1606.

Lehnertz, B., Ueda, Y., Derijck, A. A., Braunschweig, U., Perez-Burgos, L., Kubicek, S., Chen, T., Li, E., Jenuwein, T. and Peters, A. H. (2003). Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin. *Curr. Biol.* **13**, 1192-1200.

Li, L., Roy, K., Katyal, S., Sun, X., Bleoo, S. and Godbout, R. (2006). Dynamic nature of cleavage bodies and their spatial relationship to DDX1 bodies, Cajal bodies, and gems. *Mol. Biol. Cell* **17**, 1126-1140.

Li, X., Lee, Y. K., Jeng, J. C., Yen, Y., Schultz, D. C., Shih, H. M. and Ann, D. K. (2007). Role for KAP1 serine 824 phosphorylation and sumoylation/desumoylation switch in regulating KAP1-mediated transcriptional repression. *J. Biol. Chem.* **282**, 36177-36189.

Mascle, X. H., Germain-Desprez, D., Huynh, P., Estephan, P. and Aubry, M. (2007). Sumoylation of the transcriptional intermediary factor 1beta (TIF1beta), the Co-repressor of the KRAB Multifinger proteins, is required for its transcriptional activity and is modulated by the KRAB domain. *J. Biol. Chem.* **282**, 10190-10202.

Matsuda, E., Agata, Y., Sugai, M., Katakai, T., Gonda, H. and Shimizu, A. (2001). Targeting of Kruppel-associated box-containing zinc finger proteins to centromeric heterochromatin. Implication for the gene silencing mechanisms. *J. Biol. Chem.* **276**, 14222-14229..

Matsui, Y., Zsebo, K. and Hogan, B. L. (1992). Derivation of pluripotential embryonic stem cells from murine primordial germ cells in culture. *Cell* **70**, 841-847.

Meroni, G. and ez-Roux, G. (2005). TRIM/RBCC, a novel class of 'single protein RING finger' E3 ubiquitin ligases. *BioEssays* **27**, 1147-1157.

Nielsen, A. L., Ortiz, J. A., You, J., Oulad-Abdelghani, M., Khechumian, R., Gansmuller, A., Chambon, P. and Losson, R. (1999). Interaction with members of the heterochromatin protein 1 (HP1) family and histone deacetylation are differentially involved in transcriptional silencing by members of the TIF1 family. *EMBO J.* **18**, 6385-6395.

Nikulina, K., Bodeker, M., Warren, J., Matthews, P. and Margolis, T. P. (2006). A novel Kruppel related factor consisting of only a KRAB domain is expressed in the murine trigeminal ganglion. *Biochem. Biophys. Res. Commun.* **348**, 839-849.

O'Geen, H., Squazzo, S. L., Iyengar, S., Blahnik, K., Rinn, J. L., Chang, H. Y., Green, R. and Farnham, P. J. (2007). Genome-wide analysis of KAP1 binding suggests autoregulation of KRAB-ZNFs. *PLoS Genet.* **3,** e89.

Oh, H. J., Li, Y. and Lau, Y. F. (2005). Sry associates with the heterochromatin protein 1 complex by interacting with a KRAB domain protein. *Biol. Reprod.* **72**, 407-415.

Okano, M., Bell, D. W., Haber, D. A. and Li, E. (1999). DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* **99**, 247-257.

Payen, E., Verkerk, T., Michalovich, D., Dreyer, S. D., Winterpacht, A., Lee, B., De Zeeuw, C. I., Grosveld, F. and Galjart, N. (1998). The centromeric/nucleolar chromatin protein ZFP-37 may function to specify neuronal nuclear domains. *J. Biol. Chem.* **273**, 9099-9109.

Peters, A. H., O'Carroll, D., Scherthan, H., Mechtler, K., Sauer, S., Schofer, C., Weipoltshammer, K., Pagani, M., Lachner, M., Kohlmaier, A. et al. (2001). Loss of the Suv39h histone methyltransferases impairs mammalian heterochromatin and genome stability. *Cell* **107**, 323-337.

Quimby, B. B., Yong-Gonzalez, V., Anan, T., Strunnikov, A. V. and Dasso, M. (2006). The promyelocytic leukemia protein stimulates SUMO conjugation in yeast. *Oncogene* **25**, 2999-3005.

Ravasi, T., Huber, T., Zavolan, M., Forrest, A., Gaasterland, T., Grimmond, S. and Hume, D. A. (2003). Systematic characterization of the zinc-finger-containing proteins in the mouse transcriptome. *Genome Res.* **13**, 1430-1442.

Reymond, A., Meroni, G., Fantozzi, A., Merla, G., Cairo, S., Luzi, L., Riganelli, D., Zanaria, E., Messali, S., Cainarca, S. et al. (2001). The tripartite motif family identifies cell compartments. *EMBO J.* **20**, 2140-2151.

Ryan, R. F., Schultz, D. C., Ayyanathan, K., Singh, P. B., Friedman, J. R., Fredericks, W. J. and Rauscher, F. J., 3rd (1999). KAP-1 corepressor protein interacts and colocalizes with heterochromatic and euchromatic HP1 proteins: a potential role for Kruppel-associated box-zinc finger proteins in heterochromatin-mediated gene silencing. *Mol. Cell. Biol.* **19**, 4366-4378.

Schultz, D. C., Friedman, J. R. and Rauscher, F. J., 3rd (2001). Targeting histone deacetylase complexes via KRAB-zinc finger proteins: the PHD and bromodomains of KAP-1 form a cooperative unit that recruits a novel isoform of the Mi-2alpha subunit of NuRD. *Genes Dev.* **15**, 428-443.

Schultz, D. C., Ayyanathan, K., Negorev, D., Maul, G. G. and Rauscher, F. J., 3rd (2002). SETDB1: a novel KAP-1-associated histone H3, lysine 9-specific methyltransferase that contributes to HP1-mediated silencing of euchromatic genes by KRAB zinc-finger proteins. *Genes Dev.* **16**, 919-932.

Shannon, M., Hamilton, A. T., Gordon, L., Branscomb, E. and Stubbs, L. (2003). Differential expansion of zinc-finger transcription factor loci in homologous human and mouse gene clusters. *Genome Res.* **13**, 1097-1110.

Sripathy, S. P., Stevens, J. and Schultz, D. C. (2006). The KAP1 corepressor functions to coordinate the assembly of de novo HP1-demarcated microenvironments of heterochromatin required for KRAB zinc finger protein-mediated transcriptional repression. *Mol. Cell. Biol.* **26**, 8623-8638.

Sternsdorf, T., Jensen, K. and Will, H. (1997). Evidence for covalent modification of the nuclear dot-associated proteins PML and Sp100 by PIC1/SUMO-1. *J. Cell Biol.* **139**, 1621-1634.

Sutherland, H. G., Mumford, G. K., Newton, K., Ford, L. V., Farrall, R., Dellaire, G., Caceres, J. F. and Bickmore, W. A. (2001). Large-scale identification of mammalian proteins localized to nuclear sub-compartments. *Hum. Mol. Genet.* **10**, 1995-2011.

Sutherland, H. G., Lam, Y. W., Briers, S., Lamond, A. I. and Bickmore, W. A. (2004). 3D3/lyric: a novel transmembrane protein of the endoplasmic reticulum and nuclear envelope, which is also present in the nucleolus. *Exp. Cell Res.* **294**, 94-105.

Tanaka, K., Tsumaki, N., Kozak, C. A., Matsumoto, Y., Nakatani, F., Iwamoto, Y. and Yamada, Y. (2002). A Kruppel-associated box-zinc finger protein, NT2, represses cell-type-specific promoter activity of the alpha 2(XI) collagen gene. *Mol. Cell. Biol.* **22**, 4256-4267.

Urrutia, R. (2003). KRAB-containing zinc-finger repressor proteins. *Genome Biol.* **4**, 231.

Vertegaal, A. C., Andersen, J. S., Ogg, S. C., Hay, R. T., Mann, M. and Lamond, A. I. (2006). Distinct and overlapping sets of SUMO-1 and SUMO-2 target proteins revealed by quantitative proteomics. *Mol. Cell Proteomics.* **5**, 2298-2310.

White, D. E., Negorev, D., Peng, H., Ivanov, A. V., Maul, G. G. and Rauscher, F. J., 3rd (2006). KAP1, a novel substrate for PIKK family members, colocalizes with numerous damage response factors at DNA lesions. *Cancer Res.* **66**, 11594-11599.

Zhang, R., Poustovoitov, M. V., Ye, X., Santos, H. A., Chen, W., Daganzo, S. M., Erzberger, J. P., Serebriiskii, I. G., Canutescu, A. A., Dunbrack, R. L. et al. (2005). Formation of MacroH2A-containing senescence-associated heterochromatin foci and senescence driven by ASF1a and HIRA. *Dev. Cell* **8**, 19-30.

Zhong, S., Muller, S., Ronchetti, S., Freemont, P. S., Dejean, A. and Pandolfi, P. P. (2000). Role of SUMO-1-modified PML in nuclear body formation. *Blood* **95**, 2748-2752.